[V2.0 Draft as of June 30, 2022]

DATA SECURITY SPECIFIC TO VOICE

An Executive Whitepaper (DRAFT V2.0)

Executive Summary

The Open Voice Network security whitepaper is an introductory document for executives looking to use voice technologies and voice data for commercial purposes.

What you need to know: the volume and velocity of data required to build conversational AI platforms and voice services present new, complex security risks.

This document:

Provides an overview of voice-specific security risks.
Reviews strategies to accelerate enterprise adoption and the secure use and sharing of voice data.
Defines two voice-specific security principles to earn user trust.
Suggests specific next steps to advance secure multi-agent data sharing.

This paper was developed by the Privacy and Security Work Group of the Technical Committee of the Open Voice Network. Velocity.ai, a data security specialist firm based in Atlanta, Georgia, USA, served as the Work Group's technical consultant and advisor, and directly contributed to this content. 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 .

TABLE OF CONTENTS

INTRODUCTION

Voice, data, and enterprise adoption

The rise in the enterprise use of conversational artificial intelligence (AI) technologies – with investment estimated by various analyst firms to grow at a 20-30% CAGR from 2021 forward – has led to the need for this paper (Markets and Markets, 2021).

What we describe here as voice technology is the ability of a human to speak to a computer (and for the computer not only recognize meaning but respond in kind). Voice is a rapidly growing part of conversational AI; many readers are familiar with voice applications available on "smart speakers" and smartphones. Analysts estimate that more than 130 million individuals in the United States were searching for information and digital destinations through their voices as of early 2022 (Lin, 2022).

From a data perspective, voice communications are of two parts of data: (1) the *words* that are spoken, which reveal *what* the user says and (2) the *acoustics* of the utterance, which reveal *how* the user says something. through speech-to-text software, which reveal user intent. The acoustics can be used to identify the speaker (it's a biometric identifier), and may convey for example user characteristics like gender, age, and race as well as user emotions and even health status (Kröger, Lutz, Raschke, P., 2020).

Voice is one of our single richest sources of human data.

Voice technology– like all artificial intelligence -- is also rapidly advancing. Voices can now be cloned and "synthesized;" well-recognized voices can be taken apart and re-assembled to speak new scripts and even new languages. Gamers who assume new online identities with avatars can also obtain a new, unique, human-like voice – created, uniquely, to a desired tone, range, and accent.

All this opens new worlds of data analysis, of potentially significant benefits and risks to business and society. It also opens new areas of opportunity for cybercriminals, and new data security concerns for enterprise leaders.

The purpose of this paper by the Open Voice Network is to identify opportunities, risks, and solutions in the use and protection of voice data. This document will help Chief Security Officers, Chief Executive Officers (CEOs), and their boards understand the issues before them, the choices available, and why it is vital to use voice data responsibly to create sustainable business value - for enterprises and consumers alike.

In this paper, we identify four areas of voice data security that are present with the accelerating use of enterprise conversational AI. Enterprises should not assume that current, best-practice security technologies, infrastructure, and processes are sufficient to mitigate risks inherent in this new world of voice.

Risks in Plain Sight:

- General-purpose Voice Assistants and Customer and Commercial Data Acquisition

Risks Inherent with Voice Assistant Use:

- Eavesdropping
- User Authentication

Risks Inherent with Voice Assistant Development and/or Implementation

- Adversarial Attacks
- API's Creating New Threat Surfaces

New Threat Surfaces: Risks Introduced with the Interoperable Future

- Voice as a Controlling interface for the Smart Home and Smart Spaces
- Passing of Control Between Agents

This paper is for today and tomorrow. It applies to the current use of conversational AI technologies. It is also relevant to the day – coming soon – when conversational AI is the ubiquitous interface to a worldwide voice web – worthy of user trust.

A REMARKABLE TECHNOLOGY

Simply stated, voice assistance is an artificial intelligence capability that allows human speech to be understood by a computer – and for the computer to respond with speech-recognizable sound to the human question or command.

Of all the technologies that enable voice assistance, two are noteworthy: AI-driven **natural language understanding** (NLU) enables human speech to be translated into text – and for meaning to be derived from the set and order of the words. **Natural language generation** (NLG), in turn, enables an artificial intelligence-driven textual response to be turned into a human-sounding utterance.

Beneath NLU and NLG, as with all artificial intelligence, is data -- in massive abundance.   This is the data required to train NLU algorithms on the complexities of human utterance across languages, dialects, intonation, and meaning, the data necessary to synthesize human-like voices, and the data required to drive the desired predictive, anticipatory experience by users.

The process of voice understanding, and generation is data-rich. But the human voice itself brings a second (and significant) layer of data concern.

Voice data is more than words captured by microphones and translated into text. We must also consider the *acoustic* elements of voice data. At a lay level, these might include such

elements as pitch, amplitude (volume), and frequency – which, when compared to existing data sets, can be used to

- Uniquely identify the speaker (voice is a biometric identifier)
- Identify the emotion with which words were delivered (angry, excited, melancholy)
- Provide indicators of the speaker's personality (extroverted or introverted)
- Provide early identification (or progressive analysis) of an increasingly wide range of mental and physical illnesses, including respiratory conditions, Alzheimer's, Parkinson's, and schizophrenia (voice is a biomarker).

These are some of the things that can be known (or categorized) about an individual from the human voice:

Category	**Attribute**	**Example**	**Direct or inferred**
Identity	Name and related identity data	Annika	direct
Intent	Desired knowledge or action	*Buy bread*	inferred
Physical Characteristic	general identity	female	inferred
Physical Characteristic	age bracket	20's - 30's	inferred
Physical Characteristic	height, weight	tall, thin	inferred
Physical Characteristic	upper body strength	*moderate*	inferred
Ethnicity	language, dialect	German	inferred
Personality type	extrovert, introvert	introvert	inferred

Demography	geographical area	Bavarian	inferred
Demography	educational level, social class	university	inferred
Sentiment	emotion	angry	direct
Trustworthiness	intent, believability	believable	direct
Physical Condition	intoxication	no	
Physical Health	Parkinson's, Alzheimer's	no	inferred
Physical Health	fertility	no	inferred
Physical Health	infectious respiratory disease	no	inferred
Physical Health	Schizophrenia	no	inferred

Note. A table describing how voice technology can identify, analyze, and deduce sensitive personal information related to an individual's identity, physical characteristics, physical health, mental health, socioeconomic status, location, emotions, intent, and more. By Open Voice Network, 2022.

Voice, then, is an extraordinarily rich vein of persona, human data – second only, in the eyes of some experts, to that of the human genome (O'Connell, 2022).

THE FUTURE OF VOICE IN THE ENTERPRISE

Although computer-based voice recognition dates to Bell Labs' experimentation in the early 1950s (Summa Linguae, 2021), enterprises' first broad adoption of voice technology came in the 1990s with interactive voice response (IVR) systems. These allowed humans to interact with computer-operated phone systems through voice and keypad tones (Tolentino, 2015). IVR systems have been used in countless call center/customer service settings and mobile purchasing, financial services, retail ordering, travel information, and weather conditions In the ensuing years.

A new approach to voice – that of the general-purpose, consumer-centric voice assistant -- stepped into the societal spotlight in 2011, with the introduction of two Apple's Siri assistants, and again in 2014, with the launch of Amazon's Alexa assistant and Echo smart speaker. Such general-purpose assistants demonstrated substantially better voice recognition than IVR predecessors and an internet-wide ability to respond quickly to factual questions and content requests. Entrepreneurs and enterprises responded by creating the equivalent of voice applications for the general-purpose assistants; as of 2021, it was estimated that there were more than 300,000 general-purpose applications (often known as skills, actions, and capsules) worldwide (Statista, 2021).

General-purpose assistants (which were soon followed by numerous competitors worldwide) won double-digit year-on-year consumer adoption in the years 2015-2019. As of February 2021, general-purpose voice assistants were available in the United States on roughly 212 million smartphones, 127 million automobiles, and 88 million smart speakers.

However: despite the market presence, enterprise interest and investment in general-purpose voice assistance has stalled, even faded over the past years. The reasons are many, but three are key:

- Initial implementations of applications inside general-purpose voice platforms – aimed, often, at creating consumer connections or driving transactions – have often failed to produce positive returns.
- Expanding call center and customer service voice capabilities – greater recognition, more languages, textual and acoustic analysis – produced sustainable and tangible value, especially in the labor-stressed COVID-19 pandemic era.
- Senior decision-makers balked at the data-sharing requirements of the general-purpose platforms (see below.)

While enterprise investment in general-purpose voice assistants has stalled, overall enterprise interest and investment in conversational AI (especially voice) has not. In fact, there is today rapid growth in enterprise conversational AI investment across multiple vertical industries. An October 2021 study forecast a 21.8% CAGR for 2021-2026 (Markets and Markets, 2021), headlined by customer support services (the development of custom enterprise voice assistants embedded into websites, mobile apps, and products); the use of voice technology for operational efficiencies (transcription, data entry, and task management), and the use of voice interfaces in smart factories and offices.

What was the enterprise IVR is now becoming the enterprise voice assistant – AI-smart, of value to consumers, employees, and partners, able to respond accurately to millions of queries, and multi-modal (usable with and through screens of choice.) And what is now becoming the enterprise voice assistant will, in time, become part of the worldwide voice web; accessible, thanks to now-in-development industry standards, through any voice assistant on any device.

DATA SECURITY FOR VOICE IS NO SMALL TOPIC

The advent of any enterprise technology brings with it appropriate and immediate questions of data security. As it should.

In the year 2021, according to the most recent Identify Theft Resource Center (ITRC) Data Breach Annual Report (Identity Theft Resource Center, 2022), there were more data compromises reported in the United States than in any year since the first state data breach notice law became effective in 2003.

At the core of a data security strategy for voice are the best practices in people, process, and technology enunciated by global standards bodies.

Global Security Frameworks

We acknowledge these and other leading frameworks for the establishment of foundational enterprise data security:

International Telecommunications Union (ITU)

The International Telecommunications Union is a member organization of the United Nations family whose mandate is to "ensure networks and technologies seamlessly interconnect and strive to improve access to information and communication technology to underserved communities worldwide". This is accomplished through policy and regulatory activities and setting global standards and best practices.

The ITU defines a Security framework for voice-over-long-term-evolution (VoLTE) network operation. Recommendation ITU-T X.1041 analyses security threats encountered by the VoLTE network and recommends countermeasures for telecommunication operators to ensure the secure operation (ITU, 2020). These guidelines are particularly relevant for voice assistants that are operated via mobile devices (e.g., smart phones, tablets, etc.)

National Institute of Standards and Testing (NIST)

NIST supports and defines a complex number of standards frameworks to keep up with the vast and diverse innovations. One framework, the Cybersecurity Framework, is most relevant. In particular, the framework is applicable in these critical areas – 1) data protection (data-at-rest, data-in-transit, and data leaks) and

2) identity management (authentication and access control) (NIST, 2018). The NIST frameworks are technology-neutral and scalable to evolve with technology by design. They serve as a general-purpose tool to enable economies of scale. However, with this flexibility, there is no specific focus on security concerns unique to voice technology.

International Standards Organization (ISO) / International Electrotechnical Commission (IEC) Joint Task Force 1 (JTC1)

The mission of the ISO/IEC JTC1 is to develop worldwide information and communication technology standards for business and consumer applications. The standards address each pillar of security - people, processes, and technology. General standards ISO/IEC 27001 and ISO/IEC 27002 cover information and data management within an enterprise -- additionally ISO 27701:2019 added important guidance on securing personally identifiable information. Specific to voice technology, ISO/IEC 19794-13:2018 outlines the security requirements for storing, recording, and transmitting voice data. It defines a data interchange format to support a wide variety of speaker identification and verification (SIV) applications with minimal assumptions regarding voice data capture conditions or collection environment (ISO/IEC, 2018). This gap -- lack of voice data capture conditions and environment -- does not capture the identified voice-specific security risks. The guidelines are sufficiently general to support the broadest definition of speaker recognition applications and technical approaches.

Open Web Application Security Project (OWASP)

OWASP provides tools and resources to technologists to secure the web. The seminal deliverable is a Top 10 document outlining the ten most critical security concerns for web application security. The most recent update to the list was released in 2021 (OWASP, 2021). Although the security concerns apply broadly to voice technology, there are four items on the list, while not voice-specific, are relevant to the security of voice data: 1) A02: Broken Authentication/Eavesdropping, which can lead to hacking of message requests and corresponding response; 2) A03: Injection risks exists when requests are intercepted or modified; 3) A04: Insecure Design overs risks related to design and architectural flaws; 4) A07: Identification and Authentication Failures apply to missing or ineffective multi-factor authentication and session identifier most relevant in interoperability scenarios.

VOICE-SPECIFIC SECURITY RISKS

However, and why this paper is necessary: ***voice technology brings with it several unique risks and new threat surfaces that require C-suite attention.***

We see four major risk areas:

Risks in Plain Sight:
- General-purpose Voice Assistants and Customer and Commercial Data Acquisition

Risks Inherent with Voice Assistant Use:
- Eavesdropping
- Fraud Through the Use of Synthetic Voice

Risks Inherent with Voice Assistant Development and/or Implementation
- Adversarial Attacks
- API's Creating New Threat Surfaces

New Threat Surfaces: Risks Introduced with the Interoperable Future
- Voice as a controlling interface for the Smart Home and Smart Spaces
- The Passing of Control Between Agents

Risks in Plain Sight

The Data Giveaway

These are risks that hide in plain sight. And are not the result of cybercrime.

At present, many big tech general-purpose voice platforms – as part of their standard operating agreements – are entitled to ownership of all voice (text and acoustic) data that flows through the platform. An enterprise that creates and operates an application on such a platform could expose consumer preferences and sentiment and competitive commercial information (products, pricing, value propositions) to operators of the voice platform (Open Voice Network, 2021).

Voice platforms may choose to do this to provide a seamless experience to their users, such as the ability of a voice assistant to shift topics or move from an explicit to an implicit request. Platforms also do this to bolster databases for training and inference to scale the growing expanse of dialects, words, and user queries.

However, decision-makers must be aware that the implementation of an enterprise application on a general-purpose voice assistant platform brings with it a data security question. And it's not that data is at greater risk; *it's simply that it's being given away – and claimed by others – in plain sight.*

Risks Inherent in Voice Assistant Use

Eavesdropping

Eavesdropping is the unethical and unlawful act of listening to users' private conversations without explicit consent from the data owner. Most general-purpose voice assistants do not have biometric authentication capabilities built-in to the smart device. Thus, anyone can access a user's data by simply uttering a "wake word" (EDPB, 2021). A simulated conversational AI eavesdropping on ongoing talk demonstrated that a popular deep learning-based system can reliably predict if a speaker is "young," "old," "female," or "male" (age=99% gender=82%) based on what they say in around 30 seconds. The results exemplify how powerful current big data language models are in data-driven predictions of personal information based on how people talk, even when listening only for a short time (Liesenfeld et al., 2021).

To reduce the risk of a data breach at the device level, designers and developers should provide secure state-of-the-art authentication to users (EDPB, 2021). For example, voice biometric authentication capabilities can be integrated with smart devices which would only respond to the voice of the data owner. In addition, enterprises need to implement security processes that verify user consent has been given before granting anyone access to voice data (Yang et al., 2021).

In addition to user-level threats, data labeling and annotation during the model training process and in production is a risk. Data labeling and annotation processes involve human review of voice recordings and associated data raises concern due to the sensitive nature of the data. Because this process is often subcontracted out, it is essential that adequate security measures are put in place (EDPB, 2021). Enterprises need to invest in automated data preparation, annotation, labeling, and auditing tools to mitigate this risk. Enterprises need to require privileged access to audio datasets. In addition to voice anonymization technologies, background canceling technologies, homomorphic encryption, federated learning, as well as automatic detection and removal of sensitive data from text (e.g. names, addresses, credit card numbers) may help to reduce or remove biometric information from voice recordings as well as text data (Backstrom et al., 2020).

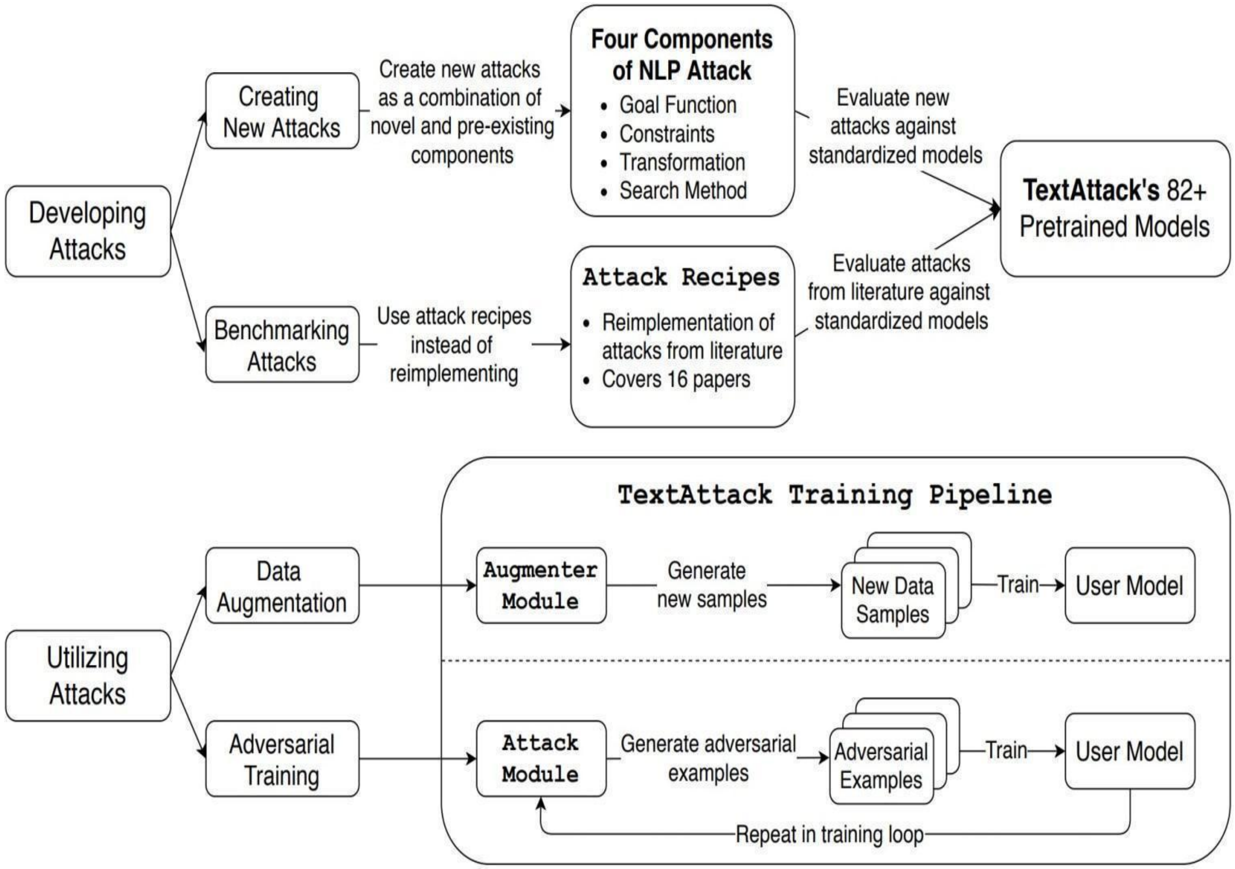
Fraud Through the Use of Synthetic Voice

Voice fraud through synthetic voice presents significant new risks to enterprises and consumers. Advances in conversational AI have introduced a new wave of voice synthesis technology, capable of producing audio that sounds like a human spoke it and with the possibility of impersonating someone's identity (Lee, 2019). Synthetic voice tools in the wrong hands will enable a range of powerful attacks against conversational AI platforms and voice services. AI-powered audio manipulation tools produce voice-deep fakes.

Conversational AI-generated deep fake voices can fool both humans and voice assistants. Simple Waveform catenation is one form of voice-synthesis technology which works by taking a person's voice and breaking it down into syllables or short sounds (Avery, 2021). Recent advances on speech synthesis increasing leverage deep learning techniques which are capable of leveraging large amounts of training data to achieve higher quality results and better performance (Ning, 2019). Cybercriminals have the capabilities today to generate data breaches and identity theft at a global scale.

For example, researchers at the University of Chicago's Security, Algorithms, Networking and Data (SAND) Lab evaluated deep fake voice synthesis programs (Avery, 2021). The researchers used tools available on the open-source developer community site GitHub to see if they could unlock voice-recognition security on enterprise conversational AI platforms. The Deep Learning and AI toolkit, known as Speaker Verification to Multi-speaker Text-To-Speech Synthesis (SV2TTS), only needed five (5) seconds of audio data to clone a human voice. The researchers demonstrated they could use SV2TTS to create voice-deep fake content at scale.

Below is an illustration of how SVTTS works:


Developing Attacks
Creating New Attacks
Create new attacks as a combination of novel and pre-existing components
Four Components of NLP Attack
• Goal Function
• Constraints
• Transformation
• Search Method
Evaluate new attacks against standardized models
TextAttack's 82+ Pretrained Models
Benchmarking Attacks
Use attack recipes instead of reimplementing
Attack Recipes
• Reimplementation of attacks from literature
• Covers 16 papers
Evaluate attacks from literature against standardized models
Utilizing Attacks
Data Augmentation
Adversarial Training
TextAttack Training Pipeline
Augmenter Module
Generate new samples
New Data Samples
Train
User Model
Attack Module
Generate adversarial examples
Adversarial Examples
Train
User Model
Repeat in training loop

(Note: SOURCE – University of Chicago SAND Lab)

Mitigating risks specific to the use of synthetic voice data

When creating a synthetic voice, among the potential ethical, economic, reputational, social, and legal harms to guard against include (Open Voice Network, 2022):

- Data breaches and identity theft.
- Unauthorized access to high-security clearance areas.
- Spreading misleading information through false content, particularly when it leads listeners to act on harmful information the speaker did not intend.

Below are recommendations on how to reduce the risks of voice deepfakes (Open Voice Network, 2022):

- Implement user identification, authentication, attestation, and authorization security processes. This should not be limited to voice identification per GDPR regulations.
- Implement synthetic voice authenticity marking technologies.
- Implement synthetic voice user recognition-detection technologies.
- Where technically possible, implement systems which do not require users to register to maintain a higher degree of anonymity.

Risks Inherent in Enterprise Voice Assistant Development and Use

Adversarial attacks on training data and algorithms

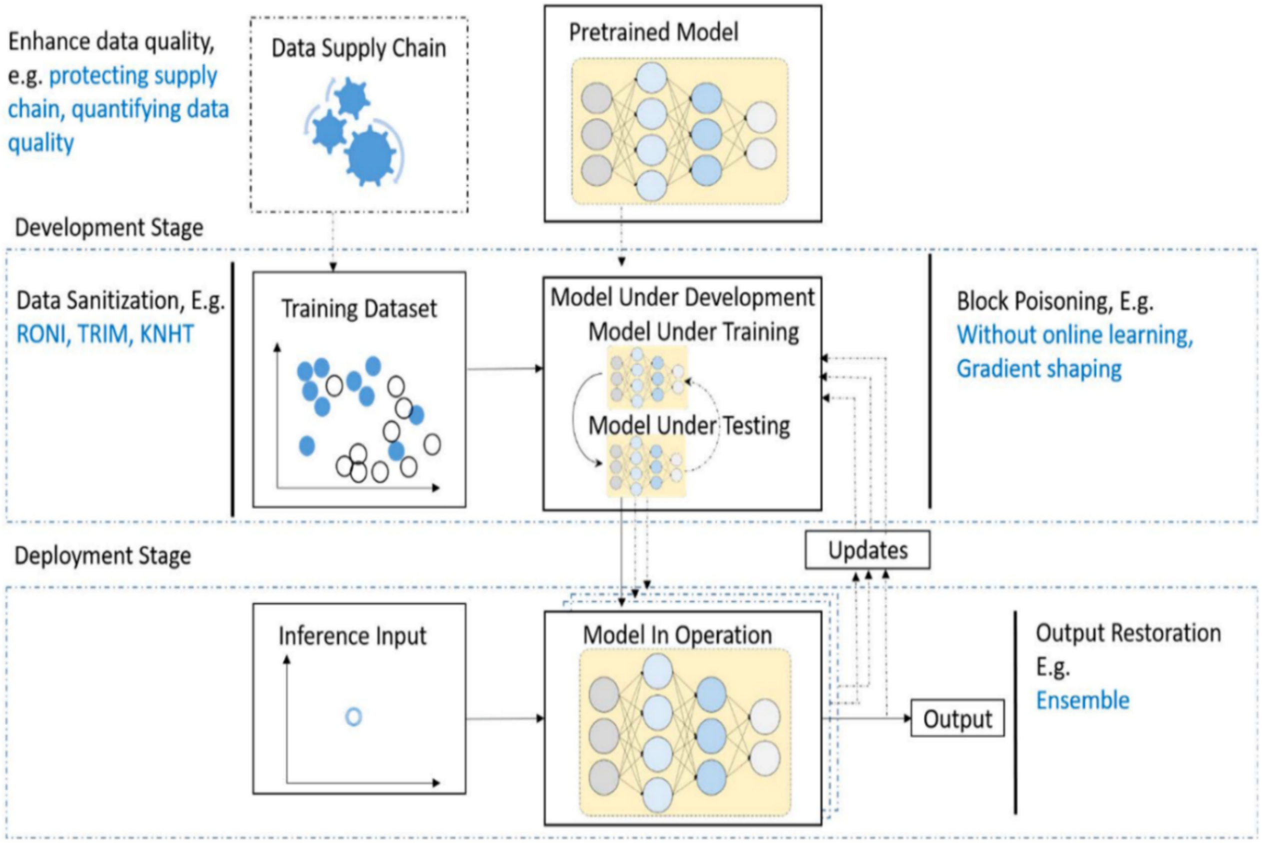
Cybercriminals use adversarial attacks to corrupt a conversational AI model's training data and algorithms. Adversarial attacks against open-source language models are an active security research area (e.g., BERT, GENIE, GPT-3, CLIP, Codex, etc.) (Fernick & Fernick, 2022; Jin, 2020).

Open-source language models use comprehensive data sets from multiple sources at scale (e.g. – audio, video, images, etc.). Experts from around the world contribute data sets to create open-source language models. Unfortunately, these open data sets used to develop language models are also available to cybercriminals who use these to build adversarial AI models.

For example, Deep Learning-based Text Understanding (DLTU) is the backbone technology behind various voice services, including question answering, machine translation, and text classification (Li et al, 2018). Despite its tremendous popularity, the security vulnerabilities of DLTU are still largely unknown, which is highly concerning given its increasing use in security-sensitive applications such as sentiment analysis and toxic content detection. According to Velocity AI, DLTU is inherently vulnerable to adversarial text attacks, in which maliciously crafted texts trigger target DLTU systems and services. Increased security innovation, specific to conversational AI, is needed to develop defense mechanisms to mitigate such attacks.

According to global market research firm Gartner, 30% of AI attacks by 2022 will involve data poisoning (Kumar & Johnson, 2021).

The diagram below shows a visual depiction of model data poisoning during model testing and training:


Enhance data quality, e.g. protecting supply chain, quantifying data quality
Data Supply Chain
Pretrained Model
Development Stage
Data Sanitization, E.g. RONI, TRIM, KNHT
Training Dataset
Model Under Development
Model Under Training
Model Under Testing
Block Poisoning, E.g. Without online learning, Gradient shaping
Deployment Stage
Updates
Inference Input
Model In Operation
Output
Output Restoration E.g. Ensemble

API's creating new risk surfaces

According to the Open Web Application Security Project, or OWASP, the most common API-related vulnerability is BOLA – broken object-level authorization, which occurs when an application checks to see if a user has authorized access privileges but doesn't verify that he or she has the correct privileges (Shkedy, 2021). Whether an application targets consumers, employees, partners or otherwise, the client-side of an application (e.g., a mobile app, a web app) interacts with the server-side of an application via an Application Programming Interface (API). Simply put, APIs make it easy for a developer to create a client-side app. Microservice architectures are also made possible by APIs (Aloufi et al., 2021).

According to new research (*Salt Security Discovers Critical API Security Vulnerability That Would Have Enabled Administrative Account Takeover on FinTech Platform Serving Hundreds of Banks*, 2022), 95% of organizations experienced an API security incident in the past 12 months. Additional research showed significant growth (681%) of malicious API traffic in the same period. APIs effectively serve as translators that allow communication and interaction between various devices and/or apps (Shkedy, 2021). A 2020 Akamai report found that nearly 20% of credential abuse attacks between 2017 and 2019 were launched against API endpoints (Barth, 2021b).

Examples of API hacker attacks have been documented:

- A skilled hacker needed only a user's message ID number to send an API request to control his or her vehicle, and then actually approve said request on behalf of the victimized vehicle operator (Shkedy, 2021). The hacker noted that the car apps did not properly use a process known as certificate pinning, which prevents man-in-the-middle attacks against APIs. (Barth, 2021b).
- The speech recognition feature of Google Chrome experienced an open-source glitch that any expert intruder could easily exploit to gain access to a computer's microphone. Most voice recognition websites work over the secure 'https' network, which means Google Chrome will not ask you every time some website asks for access to your computer's microphone. This means that the intruder at the other end of the pipe can easily sneak into the network and eavesdrop (Saleem, 2014).
- A remote code execution vulnerability exists when the Microsoft Speech API (SAPI) handles text-to-speech (TTS) input improperly. The vulnerability could corrupt memory to enable an attacker to execute arbitrary code in the current user's context (Microsoft, 2019).

With smart homes, digital assistants, chatbots, and more, Google claims that we live in an age of voice experiences. There's a pressing need to enable voice experiences as quickly as other more traditional interfaces, such as mobile apps and websites. As a result, they expect chat and voice platforms with conversational APIs will release them to the rest of the world to expand their reach (Vigliarolo, 2022b).

However, most companies are ill-prepared to defend against an API attack because traditional security tools such as web application firewalls (WAFs) and API gateways cannot detect API manipulation. The consequences can be severe, both monetary and reputational damage (*Salt Security Discovers Critical API Security Vulnerability That Would Have Enabled Administrative Account Takeover on FinTech Platform Serving Hundreds of Banks*, 2022).

Risks Introduced by a Future of Voice Interoperability

Voice technology is soon to enter an age of system-to-system interoperability – one that, from a data security perspective, is likely to open new threat surfaces.

The value of standards-based interoperability has been well documented, and through the generations; markets grow, innovation accelerates, and development and implementation costs drop, with benefits accruing to all parties in an ecosystem (*Standards Boost Business: Value of Standards*, 2022). For participating enterprises, standards-based interoperability allows "build once, use many" investments; resources can be applied to brand and experience differentiation instead of building and managing multiple proprietary instantiations of the same functionality (O'Connor, 2017).

As we look ahead to the NLP-NLG future, we anticipate that enterprise interests and consumer demand will soon lead to the exchange of textual and acoustic data, as well as "context" (data regarding prior use and stated intent) between different voice systems. Controls for privacy and security will also be passed. Envision a chain of dialogues; a weak security link could put the entire chain at risk.

Three interoperability scenarios are now before security specialists and researchers:

- NLP-NLG as an interoperable interface for smart home and smart space IoT systems
- NLP-NLG as the interface for and to interoperable voice assistants and voice-enabled websites
- NLP-NLG as the interface to a personal AI "agent" capable of independent decision-making and actions with other AI agents on behalf of the human user.

Voice as a controlling interface for the smart home and smart spaces

Internet of things (IoT) is the descriptive term for a system of interrelated computing devices, mechanical and digital machines, objects, animals, or people that are provided with unique identifiers and the ability to capture and transfer data through a digital network without human intervention (Gillis, 2022).

A “thing” in IoT could be an implanted heart monitor in a human, an air pressure sensor in an automobile’s tire, or a set of temperature and humidity sensors within a home or factory -- all which can be assigned an Internet Protocol (IP) address and are able to send data through a public or private network (Gillis, 2022).

Given the screen-free, hands-free convenience of voice – and the value of multi-language understanding -- it is not surprising that NLP-NLG technologies are becoming an interface of choice for IoT implementations, especially in the consumer “smart home” and in enterprise “smart spaces.”

A 2021 European Commission white paper (European Commission, 2021) noted both the rapid market growth forecast for consumer IoT and within that growth, the important role of voice interfaces. Overall consumer IoT expenditures are predicted to more than double through this decade, from roughly $125.4B (USD) in 2019 to $480.2B (USD) by 2030. Voice assistants – and specifically, general-purpose Big Tech voice assistants – are expected to become the primary interface to smart devices and consumer IoT services.

The EU paper also spoke to the need for interoperability within the consumer IoT ecosystem – and not only among smart devices, but among voice assistants. According to the white paper, general-purpose voice assistants serve as the consumer “entry point” to smart devices; the lack of interoperability among the leading general-purpose voice assistants limits market access and opportunity for smart device and services providers, and overall growth of the consumer IoT market (European Commission, 2021).

Voice is also expected to serve as a primary interface for the operation of enterprise IoT, often referred to as “smart spaces.” According to a March 2022 forecast from the firm IoT Analytics, global enterprise spending on IoT implementations was expected to increase from 2022 to 2027 at a 22% compound annual growth rate, from roughly $194B (USD) to $525B (USD) (Wegner, 2022).

The value propositions of enterprise IoT are many and include automated facility energy management, and the use of sensors (heat, friction, humidity, etc.) for pre-emptive manufacturing maintenance and the optimal use of resources (water, fertilizer, etc.) in agriculture.

Passing of control between voice agents

Interoperability between voice assistants and voice agents brings with it the passing of data (textual, acoustic, and contextual) and control (privacy and security) from the NLP-NLG infrastructure one to another. The Open Voice Network is currently developing proposed standards for the seamless passing of data and controls between NLP-NLG systems.

However, and as security experts know: the passing of data beyond firewalls and to systems of different parentage is a reason for security concern. In addition, the potential of interoperability between voice *agents* is the reason for new data security research.

Both voice assistants and voice agents:

- provide requested information using the NLP-NLG infrastructure of a conversational platform
- are perceived by users to be a single conversational actor
- provide continuity of knowledge and persona
- can fulfill complex user intents through delegation and receipt of control
- have an addressable name, such as “Alexa,” or “Siri,” “Erica, or “Magenta.” (Open Voice Network Technical Committee, 2022).

In addition, however, a voice agent can operate *independently on behalf of the user*.

The difference may be understood this way:

- a voice assistant can do only what it is told to do. It can provide fast and accurate access to all knowledge stored within its walls.
- A voice agent will, over time and thanks to the learning powers of artificial intelligence, know the interests and preferences of its primary user – and act upon those on behalf of the user. In the years ahead, agents will likely manage personal calendars (make appointments, book tables and travel), shop (perhaps in auctions involving the agents of retailers worldwide), obtain reference information and make introductions.

This leads us to two scenarios – one immediately ahead of us, and one that is now coming into focus.

- Immediately before us, **a consumer – using a general-purpose voice *assistant or agent* – will want to connect beyond general-purpose applications to full immersion in an independent, brand-based *assistant.*** Through the general-purpose voice agent or assistant, the consumer will connect with and fully enter a favorite brand's voice- or multi-modal assistant experiences. Potential reasons to do so are many: consumers may want to explore a broader range of products and services, enjoy preferential status with the brand, or seek a higher level of personal privacy than the app would allow. From the other direction, an enterprise may wish to step beyond its voice assistant firewall and communicate a message (a reminder, a birthday wish, a special offer, health-related services) to a consumer who uses a general-purpose voice assistant.

- Now coming into focus, **a consumer – using a general-purpose or independent, personal *agent* with a voice interface – will ask the agent to complete a task.** If it's grocery shopping, the personal agent may contact the agents of multiple retailers and determine, in milliseconds, the optimum combination of product, price, and delivery times by provider in accord with the consumer's preferences. If it's travel, the personal agent may connect with an airline, hotel, and event ticket opportunities to develop optional itineraries according to the needs and wants of the consumer**.** Data may not be exposed (and exchanged) sequentially in this scenario – but to multiple destinations simultaneously.

It's a reason for this paper. Continued research by the Open Voice Network to bring security protocols (and standards) to the world of new threat surfaces created by interoperability.

OPPORTUNITY

Earn user trust by creating secure voice services

The growth of the conversational AI market presents a unique opportunity for enterprises to earn user trust by implementing secure voice services that protect user data. In contrast, due to the alarming increase in data breaches worldwide, many people have lost confidence in any enterprise's ability to protect their data. As a result, many consumers are reluctant to share their data with businesses. Solving security problems with new standards, principles, and data infrastructure that protects voice data is critical to the success of the global voice ecosystem.

According to global consultancy Deloitte, conversational AI solutions—including conversational agents, chatbots, and voice assistants—have become extraordinarily popular with accelerated adoption in the enterprise due to the COVID-19 pandemic. Deloitte's market research reveals that "data from various conversational AI vendors showed that the volume of interactions handled by conversational agents increased by as much as 250% in multiple industries". Implementing better security will undoubtedly accelerate enterprise adoption of conversational AI technologies (Comes et al., 2021).

SOLUTION

There are many stakeholders in voice-specific security standards. The dramatic increase in data breaches worldwide proves that new security standards and technologies are needed to secure conversational AI.

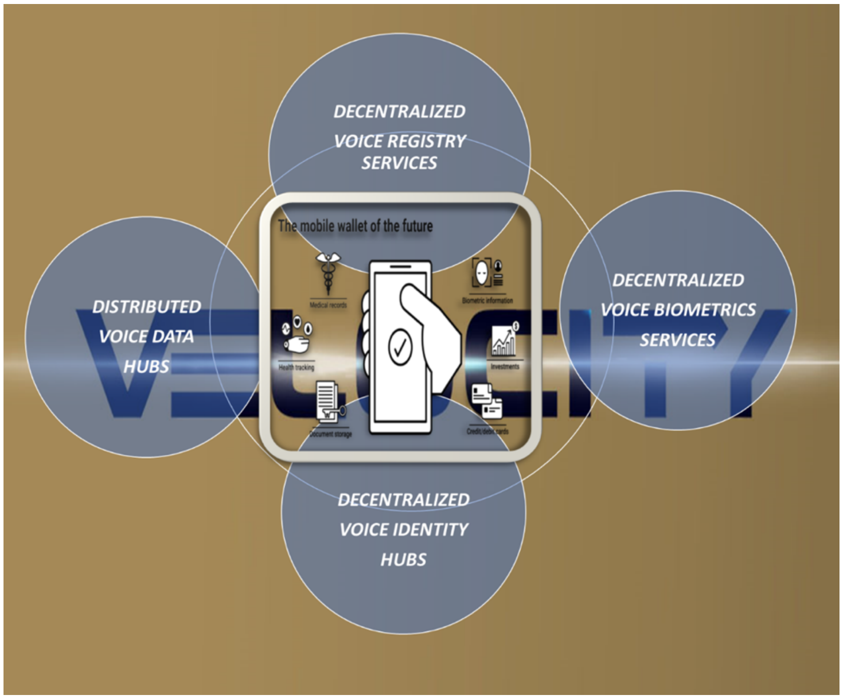
The Open Voice Network asserts (Open Voice Network, 2022):

- Data protection and secure identity -- Businesses own their customer data. Consumers own their identity (ID) data. Conversational AI platforms should have on-device security intelligence capabilities that protect user data by default. Enterprises should only deploy conversational AI platforms and voice services when they have the security capabilities to protect consumer data and identities.

- Data trustworthiness and authenticity -- Businesses should only share voice data with explicit consent from the data owner. Companies should provide a separate opt-in for conversation data analytics purposes to allow users to give explicit consent to use their data for analytics. Enterprises should balance speed to market with responsible AI practices that center on security by design. Enterprises should have security capabilities to detect an authentic human voice from a synthetic voice.

Open Voice Network Security Principles

Enterprises need to invest in new security standards to build trust in conversational AI platforms and voice services. We acknowledge and applaud the efforts of global government entities in establishing policies to address this growing concern.



(Source – VELOCITY AI – VOICE Data Security Reference Architecture)

The Open Voice Network standards proposal is an answer to the "Call to Action" for new global standards to protect the "voices" of businesses and consumers. We identify the gap, the opportunity, and the need -- and recommend the implementation of standards in two areas:

1. Secure voice identification of users and conversational AI agents
2. Secure multi-agent data sharing

We believe that improving security will accelerate enterprise adoption of AI-powered voice technologies for commercial purposes.

1. **Voice-Specific Security Standard**: Secure Voice Identification of Users and Conversational AI Agents

The European Data Protection Board ('EDPB') published its guidelines on voice assistance. The guidelines identify the most relevant compliance challenges and provide recommendations to relevant stakeholders (European Union, 2021). In particular, the guidelines note that data controllers providing voice services and their processors must consider both the General Data Protection Regulation (European Union, 2016) and the Directive on Privacy and Electronic Communications (Office des Publications Officielles des Communautés Européenne, 2002).

Specifically, the guidelines highlight the increased complexity for data controllers who provide such services in meeting GDPR transparency requirements because of the multiple users of voice assistants, the ecosystem complexities, and the specificities of the vocal interface.

Moreover, the guidelines recommend that users be informed at the earliest time possible and, at the latest, at processing time. The guidelines also recommend informing users of the purposes of personal processing data, which should accord with their expectations of the device they purchase.

Enterprises must provide a separate opt-in for conversation data analytics purposes to allow users to give explicit consent for how their data is processed (European Data Protection Board, 2021). If voice control is by default activated, the device must deactivate speech data analysis until the user gives explicit consent. In addition, the guidelines note that the current parental controls framework for voice assistance is not user-friendly and urges data controllers to invest in developing means for parents or guardians to control children's access to devices.

The United States (U.S.) Federal Trade Commission (FTC) has published guidance on how consumers can secure their voice assistants. The FTC makes the following security recommendations (Federal Trade Commission, 2020):

- "Each time you interact with it, your voice assistant records what you say. It might also do that when it thinks it's heard the wake word."
- "If you want to ensure that your smart speaker doesn't pick up sensitive information, look for settings to mute your device, so it's no longer listening."

- The FTC also recommends activating alerts that tell you when your voice assistant is actively listening. "Some voice assistant manufacturers have had employees listen to audio recordings."

The Open Voice Network identifies the significant responsibility of the consumer to overcome the lack of secure-by-design standards among platforms and independent voice agents. The complexity is compounded in multi-agent data sharing when there may not be a voice user interface to give explicit consent to the use, ownership, or sharing of voice data. **We assert the need for secure voice identification of users and conversational AI agents.**

2. Voice-Specific Security Standard: Secure Multi-Agent Data Sharing

The Open Voice Network security standards will improve interoperability – between conversational agents and risk management processes for improved customer experiences.

Secure multi-agent data sharing is a new AI paradigm for trustworthy conversational analytics. After receiving their explicit consent, AI agents share data on behalf of the user (enterprise employees or consumers). Thus, data remains securely held on the user's device (voice-controlled mobile wallet). Combining **confidential computing** and **differential privacy techniques** will protect user identities and data processing (reference following sections). This approach enables enterprises to use voice data for commercial purposes - highly secure and privacy-preserving.

- **Confidential Computing for Protecting Voice Data in Use**

  Confidential computing security infrastructures use hardware-based semiconductors (chips) called trusted encryption environments (TEE) and homomorphic encryption algorithms to protect data – thus reducing the risks to foundational language models. This methodology will reduce the chances of unauthorized access to private consumer data by using TEE hardware combined with homomorphic software to protect conversation training data sets and secure data analytics workloads (Confidential Computing Consortium, 2020).

- **Differential Privacy for Protecting Personally Identifiable Data in Use**

  Differential privacy is a security method for privately sharing information about a dataset by describing the patterns of groups within the dataset while withholding information about individuals. Important to note that existing anonymization methods are susceptible to data breaches. With a small data set containing PII, an individual's identity is reverse engineered from the original dataset. In contrast, using differential privacy algorithms - PII is not removed from a data set before performing analysis. Instead, a machine learning technique called statistical noise is inserted into the data set (Microsoft, 2018). OpenDP is a community effort to build reliable, open-source software tools for sensitive private data statistical analysis. These security tools are designed to protect personal data. According to OpenDP, differential privacy is the gold standard of privacy protection (OpenDP, n.d.).

Protecting voice data in use – is a new frontier. In the digital world where we are constantly storing, consuming, and sharing sensitive data - from credit card data to medical records, from firewall configurations to our geolocation data - protecting sensitive data in all its states is more critical than ever (Intel, n.d.). **The Open Voice Network asserts the need for secure data in use in multi-agent voice data sharing and advocates for implementing confidential computing and differential privacy as solutions**.

NEXT STEPS

This white paper is Step 1 in a framework for a complete action plan. We identify voice-specific security risks and assert the need for two security standards to protect enterprise and consumer voice data:

- Secure voice identification of users and conversational AI agents
- Secure multi-agent data sharing

We recognize that there is still more work on voice-specific data security standards. To provide enterprise decision-makers with a set of resources and tools to help navigate the security concerns, we identify two additional immediate next steps towards the development of:

- Step 2: Security Reference Implementation – identify a security use case to demonstrate recommended security architecture aligned with Open Voice Network Interoperability Standards.
- Step 3: Security Standards Certification – provide enterprise providers with tools to certify adherence to Open Voice Network security standards.

About the Open Voice Network

The Open Voice Network (OVON) is a non-profit industry association dedicated to developing 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 on 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.

Acknowledgments

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, Anna Bulakh, Oita Coleman, Ali Dalloul, Jonathan Eisenzopf, Mike Frazzini, John Iwasz, Roger Kibbe, Jim Larson, Maarten Lens-Fitzgerald, Brenda Leong, Anna Leschanowsky, Lawrence Lin, Nick Myers, Michael Novak, Sara Oliver, Dr. Birgit Poppe, Brian Owen, Shyamala Prayaga, Doug Rogers, Ed Sewell, Carl Scott, Nathan Southern, Jon Stine, and John Trammell.

Reference List

Allman-Ward, M., Sagner, & J. (2003). Essentials of managing corporate cash. Chapter 7 International Cash Management. New Jersey: JohnWiley & Sons, Inc. https://scholarworks.bridgeport.edu/xmlui/bitstream/handle/123456789/343/Sagner%20Essentials%20of%20managing%20corporate%20cash%20Chapt%2007.pdf

Aloufi, R., H., H., & Boyle, D. (2021, July). A Tandem Framework Balancing Privacy and Security for Voice User Interfaces. Imperial College London. https://arxiv.org/pdf/2107.10045.pdf

Avery, Dan. October 2021, SOURCE - DAILY MAIL UK AI-generated deepfake voices can fool both smart https://www.dailymail.co.uk/sciencetech/article-10081007/amp/AI-generated-deepfake-voices-fool-smart-assistants-humans-5-seconds-training.html

Backstrom, T., Bruggemeier, B., & Fischer, J. (2020, March 4). Privacy in Speech Interfaces. ITG News. https://www.vde.com/resource/blob/1991012/07662bec66907573ab254c3d99394ec7/itg-news-juli-oktober-2020-data.pdf

Barth, B. (2021, September 22). *Rapid proliferation of APIs opens up new security holes*. SC Media. https://www.scmagazine.com/analysis/application-security/rapid-proliferation-of-apis-opens-up-new-security-holes

Biswas, Debmalya. 2020. Towards Data Science. Privacy Risks of Chatbot Conversations. https://towardsdatascience.com/hidden-privacy-risks-of-chatbot-conversations-881dbeeb98a)

Braunlein, F., Frerichs, L., & Security Research Labs. (2019, October 20). Smart Spies: Alexa and Google Home expose users to vishing and eavesdropping. Https://Www.Srlabs.de/Bites/Smart-Spies. https://www.srlabs.de/bites/smart-spies

Center for Security and Emerging Technologies at Georgetown University - AI and the Future of Disinformation Campaigns Part 2: A Threat Model Katerina

SedovaChristine McNeillAurora Johnson Aditi Joshi Ido Wulkan December 2021)

Claypoole, T. (2021, September 28). Voice Analysis Complicates Personal Privacy. The National Law Review. https://www.natlawreview.com/article/voice-analysis-complicates-personal-privacy

Comes et al, 2021. Conversational AI. https://www2.deloitte.com/xe/en/insights/focus/signals-for-strategists/the-future-of-conversational-ai.html

Confidential Computing Consortium. (2020, October). Confidential Computing Deep Dive v1.0. Https://Confidentialcomputing.Io/Wp-Content/Uploads/Sites/85/2020/10/Confidential-Computing-Deep-Dive-White-Paper.Pdf.

CBS News. (2021, July 5). How U.S. cyber policy changed after SolarWinds. https://www.cbsnews.com/news/solarwinds-60-minutes-2021-07-04/.

Cybersecurity & Infrastructure Security Agency. (n.d.). Defining Insider Threats. https://www.cisa.gov/defining-insider-threats

Elliot, J. (2021, July 29). What You Need To Know About AI Security — Even If Your Company Isn't Using AI Yet. Forbes. https://www.forbes.com/sites/forbestechcouncil/2021/07/29/what-you-need-to-know-about-ai-security---even-if-your-company-isnt-using-ai-yet/?sh=141a64c810a0

European Commission. (2021, June 9). *Commission Staff Working Document: Preliminary Report - 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

Deepfake of Lola Flores for the new Cruzcampo ad. (2021, January 21). Tokyvideo.Com. https://www.tokyvideo.com/video/deepfake-of-lola-flores-for-the-new-cruzcampo-ad

Fathima, S. (2020, September). What is Differential Privacy? Https://Becominghuman.Ai/What-Is-Differential-Privacy-1fd7bf507049. https://becominghuman.ai/what-is-differential-privacy-1fd7bf507049

Federal Trade Commission. (2022, January 31). How to Secure Your Voice Assistance to Protect Your Privacy. https://consumer.ftc.gov/articles/how-secure-your-voice-assistant-protect-your-privacy

Fernick, J., & Fernick, V. A. P. B. J. (2022, January 1). On the malicious use of large language models like GPT-3. NCC Group Research. https://research.nccgroup.com/2021/12/31/on-the-malicious-use-of-large-language-models-like-gpt-3/

Gillis, A. S. (2022, March 4). *What is the internet of things (IoT)?* IoT Agenda. https://www.techtarget.com/iotagenda/definition/Internet-of-Things-IoT

IBM Security, 2021. Cost of Data Breach Report. https://www.ibm.com/security/data-breach

Identity Theft Resource Center. (2022). Data Breach Annual Report. Data Breach Annual Report.

Intel. (n.d.). Swiss Re Explores Further Protection of Critical Data. https://www.intel.com/content/dam/www/central-libraries/us/en/documents/swiss-re-sgx-case-study.pdf

ISO/IEC. (2018). Information technology -- Biometric data interchange formats: Voice Data. Https://Www.Iso.Org/Obp/Ui/#iso:Std:Iso-Iec:19794:-13:Ed-1:V1:En. https://www.iso.org/obp/ui/#iso:std:iso-iec:19794:-13:ed-1:v1:en

Jin, D., Zhijin, J, Zhou, J., Szolovits, P. (2020, April 8). Is BERT Really Robust? A Strong Baseline for Natural Language Attack on Text Classification and Entailment. Computer Science & Artificial Intelligence Lab, Massachusetts Institute of Technology. https://doi.org/10.1609/aaai.v34i05.6311

Kröger, J. L., Lutz, O. H., & Raschke, P. (2020). *Privacy Implications of Voice and Speech Analysis – Information Disclosure by Inference*. SpringerLink. https://link.springer.com/chapter/10.1007/978-3-030-42504-3_16

Kumar, R. S. S., & Johnson, A. (2021, March 16). Cyberattacks against machine learning systems are more common than you think. Microsoft Security Blog. https://www.microsoft.com/security/blog/2020/10/22/cyberattacks-against-machine-learning-systems-are-more-common-than-you-think/

Lazzarotti, J., & Atrakchi, M. (2020, December 10). As Voice Recognition Technology Market Surges, Organizations Face Privacy and Cybersecurity Concerns. The National Law Review. https://www.natlawreview.com/article/voice-recognition-technology-market-surges-organizations-face-privacy-and

Lee, D. (2019, May 10). "Deepfake Salvador Dali takes selfies with museum visitors". https://www.theverge.com/2019/5/10/18540953/salvador-dali-lives-deepfake-museum

Li, J., Ji, S., Du, T., Li, B., & Wang, T. (2019). TextBugger: Generating Adversarial Text Against Real-world Applications. Proceedings 2019 Network and Distributed System Security Symposium. https://doi.org/10.14722/ndss.2019.23138

Liesenfeld, A., Parti, G.. Huang, C. (2021). Deep Learning Meets Private Talk: Conversational AI Can Predict Speaker Traits by Eavesdropping for Only 30 Seconds. Mensch und Computer. Pp 547-551. [OVON research library Google Drive]

Make AI Trustworthy by Building Security Into Your Next AI Project. (2021, June 12). Gartner. https://www.gartner.com/smarterwithgartner/how-to-make-ai-trustworthy

Markets and Markets. (2021, October). Opportunities in the Conversational AI Market.

Microsoft. (2022, February 23). Data, privacy, and security for Custom Neural Voice – Azure Cognitive Services. https://docs.microsoft.com/en-us/legal/cognitive-services/speech-service/custom-neural-voice/data-privacy-security-custom-neural-voice

Microsoft. (2018). Decentralized Identity. https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE2DjfY

Microsoft. (2021). Differential Privacy for Everyone. https://drive.google.com/drive/u/2/folders/1-Snpacfbq2s_dMpHWiYSJstaP3h4LQ59

Microsoft. (2019, June 11). *Security Update Guide - Microsoft Security Response Center*. https://msrc.microsoft.com/update-guide/vulnerability/CVE-2019-0985

National Institute of Standards and Testing (NIST). (2018, April 16). Cybersecurity Framework. NIST Cybersecurity Framework. https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf

Nicholas, C., Wagner, D. (2018, March 30). Audio Adversarial Examples: Targeted Attacks on Speech-to-Text. Https://Arxiv.Org/Pdf/1801.01944.Pdf.

Ning, Y. (2019, September 27). A Review of Deep Learning Based Speech Synthesis. MDPI. https://www.mdpi.com/2076-3417/9/19/4050

O'Connell, H. Canary Speech, (March 2022). Personal conversation.

O'Connor, S. (2017, May 30). *What Is Interoperability, and Why Is it Important?* https://www.adsc.com/blog/what-is-interoperability-and-why-is-it-important

Office des Publications Officielles des Communautés Européenne. (2002). 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). Https://Eur-Lex.Europa.Eu/LexUriServ/LexUriServ.Do?uri=CELEX:32002L0058:En:HTML.

OpenDP. (n.d.). What Is Differential Privacy? (n.d.). Https://Opendp.Org/About.

Open Voice Network. (2021). "Platform Data Use and Privacy Policies: Research. Unpublished research paper.

Open Voice Network. (2022). Synthetic Voice for Content Owners and Creators. [Final website link to document]

Open Voice Network. 2022. Ethical Guidelines for Voice Experiences v1.0. https://openvoicenetwork.org/documents/ovn_ethical_guidlines_voice_experiences.pdf

Open Voice Network Technical Committee, (2022 February 23)

Open Web Application Security Project. (2021). OWASP Top 10. Https://Owasp.Org/Top10/. https://owasp.org/Top10/

Rosenzweig, P. (2021, August 31). Enterprise Cybersecurity Measurement. Lawfare. https://www.lawfareblog.com/enterprise-cybersecurity-measurement

Saleem, F. (2014, February 19). *Serious security vulnerability in Chrome speech recognition discovered*. Innov8tiv - Blacks In Technology in USA, UK, Caribbean Islands & Africa. https://innov8tiv.com/serious-security-vulnerability-chrome-speech-recognition-discovered/

Salt Security Discovers Critical API Security Vulnerability That Would Have Enabled Administrative Account Takeover on FinTech Platform Serving Hundreds of Banks. (2022, April 7). https://www.prnewswire.com/news-releases/salt-security-discovers-critical-api-security-vulnerability-that-would-have-enabled-administrative-account-takeover-on-fintech-platform-serving-hundreds-of-banks-301519736.html

Sedova, K., McNeill, C., Johnson, A., Joshi, A., Wulkan, I. Center for Security and Emerging Technology. AI and the Future of Disinformation Campaigns Part 1. https://cset.georgetown.edu/wp-content/uploads/CSET-AI-and-the-Future-of-Disinformation-Campaigns.pdf

Standards Boost Business: Value of Standards. (2022, May). https://www.standardsboostbusiness.org/value_standards.aspx

Summa Linguae. (2022, March 7). Speech Recognition Software: Past, Present, and Future. https://summalinguae.com/language-technology/speech-recognition-software-history-future/

Swiss Re explores further protection of critical data. (n.d.). https://www.intel.com/content/dam/www/central-libraries/us/en/documents/swiss-re-sgx-case-study.pdf

Tolentino, J. (2015). Enhancing customer engagement with interactive voice response. The Next Web.

Turow, J. (2021). The Voice Catchers: How Marketers Listen In to Exploit Your Feelings, Your Privacy, and Your Wallet. Yale University Press.

The White House. (2021, May 13). Executive Order on Improving the Nation’s Cybersecurity. The White House. https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/

Vigliarolo, B. (2022b, April 26). *Google Cloud sees storm brewing over API security*. https://www.theregister.com/2022/04/26/google_cloud_api/

Wegner, P. (2022, March 30). *Global IoT market size grew 22% in 2021 — these 16 factors affect the growth trajectory to 2027*. IoT Analytics. https://iot-analytics.com/iot-market-size/

Wiggers, K. (2021, May 28). Adversarial attacks in machine learning: What they are and how to stop them. VentureBeat. https://venturebeat.com/2021/05/29/adversarial-attacks-in-machine-learning-what-they-are-and-how-to-stop-them/

XSTR-SUSS - Successful use of security standards (2nd edition). (2020). ITU. https://www.itu.int/pub/T-TUT-SEC-2020-1

Yang, W., Wang, S., Sahri, N. M., Karie, N. M., Ahmed, M., & Valli, C. (2021). Biometrics for Internet-of-Things Security: A Review. Sensors (Basel, Switzerland), 21(18), 6163. https://doi.org/10.3390/s21186163

Licensing and Attribution