

Executive Summary

Sponsor : Courtney Yang (Meta)

Advisor : Lorrie Cranor (Carnegie Mellon University)

Team Members : Ruiyang Liu, Ashutosh Sahu, Prahaladh Chandrahasan (Carnegie Mellon University)

Motivation

Voice assistants have become ubiquitous in American households, with nearly half of the population utilizing these "always-listening" devices for hands-free convenience. However, this adoption is accompanied by privacy concerns, ranging from phantom activations to third-party data sharing. Moreover, users are currently forced to resort to companion applications to execute their privacy preferences, creating a "modality mismatch" between daily interaction and privacy management. This disconnect, coupled with the convoluted nature of privacy policies and obscure privacy choices, results in incomplete mental models and ultimately erodes user trust in their voice assistants. This research is motivated by the opportunity to bridge this gap by exploring whether delivering privacy notices and enabling controls directly through the voice channel could enhance transparency, reduce friction, and build greater trust between users and their devices.

The three main research questions of the study are listed below

- What privacy concerns exist with the current user experience of voice assistants' privacy notices and choices?
- What are users' perspectives, preferences, and suggestions for receiving privacy notices and making privacy choices for voice assistants through the voice channel?
- Can a designed voice user interface for privacy notices in voice assistants increase the transparency of privacy notices and choices for users and, in turn, build more trust with voice assistants?

A literature review regarding current user attitudes and design recommendations , followed by a competitive analysis of the top three voice assistants in the USA (Alexa, Google Home, and Apple HomePod), revealed a systematic exclusion of privacy management from voice interfaces. These devices consistently failed to answer privacy queries or redirected users to external apps. This finding underscored the urgent need for native voice-based privacy controls, which subsequently became the central design philosophy of our prototype.

Study Design

To answer the research questions, we conducted a semi-structured interview study with 11 voice assistant users, combining open-ended background inquiries with a structured evaluation of a novel VUI prototype. The study centered on "Nexa," a functional voice assistant prototype powered by a Large Language Model (Claude Haiku 4.5) designed to deliver privacy notices and execute choices through natural conversation. The evaluation utilized a counterbalanced within-subjects design where participants completed three standardized privacy tasks—understanding data collection, opting out of interest-based advertising, and deleting voice recordings—using both the voice prototype and a simulated graphical user interface (GUI) modeled after the Alexa app. This dual-modality approach allowed for a direct comparison of user sentiment, measuring perceived transparency, control, and information completeness across both interfaces. The study also included detailed background questions regarding participants' current privacy behaviors, such as policy reading habits and shared device usage, to contextualize their reactions to the prototype.

Results, Limitations and Recommendations

The study revealed that while the vast majority of participants expressed a strong willingness to use voice interfaces for privacy tasks due to their hands-free convenience and efficiency for quick inquiries, preferences were highly context-dependent and favored a hybrid approach rather than a complete replacement of graphical interfaces. Participants generally preferred the VUI for initial discovery and low-effort execution but retained a strong preference for the GUI when performing high-stakes actions like data deletion or opting out of tracking, citing the critical need for visual confirmation to verify that their choices were actually implemented. Limitations of this work include the small, predominantly tech-savvy sample size and the inherent variability of the LLM backend, which occasionally produced hallucinations despite strict prompting instructions. Consequently, we recommend that designers treat voice interactions as a complementary channel that proactively surfaces concise privacy summaries and leverages Large Language Models for multi-turn explanations, while still allowing users to visually verify complex privacy decisions.