# Enhancing Transparency in AI Voice Assistants

Ruyiang Liu
School of Computer Science
Carnegie Mellon University
ruiyang3@cs.cmu.edu

Ashutosh Sahu
School of Computer Science
Carnegie Mellon University
ashau2@cs.cmu.edu

Prahaladh Chandrahasan
School of Computer Science
Carnegie Mellon University
prahalac@cs.cmu.edu

December 2025

# 1 Executive Summary

## 1.1 Motivation and Research Questions

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 user uncertainty regarding privacy practices associated with their voice assistants; this lack of clarity ultimately erodes user trust in their devices. 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 assistants.

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 voice user interface for privacy notices 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.

## 1.2 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 voice user interface (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

## 1.3 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. Their preferences were highly context-dependent, and they 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 instructions. Consequently, we recommend that designers treat voice interactions as a complementary channel

that proactively surfaces concise privacy summaries and leverage Large Language Models for multi-turn explanations, while still allowing users to visually verify their privacy preferences.

## 2 Introduction

Imagine having a conversation at home about planning a vacation to Hawaii. You never explicitly asked your voice assistant for travel recommendations, yet the next time you browse online, you notice advertisements for Hawaiian resorts appearing on your shopping feeds. This scenario illustrates a common privacy concern with voice assistants called phantom activation, where devices inadvertently record conversations without clear user awareness or intent, and in this case, your voice recordings are used to for targeted advertising[3]. While such accidental triggers represent one category of privacy risk, they reflect a broader set of concerns users face when living with always-listening devices in their homes.

Voice assistants have become increasingly prevalent in American households. According to Pew Research Center, 46% of Americans use digital voice assistants, with 8% using voice assistants on standalone devices [12]. Among the reasons people adopt these technologies, 55% cite the ability to interact with devices hands-free as a primary motivation [12]. This hands-free interaction paradigm represents a shift in how users engage with technology. Rather than navigating visual interfaces through touch or mouse input, users speak naturally to accomplish tasks ranging from playing music and setting timers to controlling smart home devices and retrieving information. Voice has become the primary modality through which users interact with these assistants, establishing user expectations that voice-based interaction should be sufficient for all assistant-related tasks.

Despite widespread adoption, voice assistant users harbor significant privacy concerns about these always-listening devices. Research has documented user worries about when devices are actively recording, how audio data is processed and stored, whether conversations are reviewed by human employees, and how collected information might be shared with third parties or used for targeted advertising [7][14][10]. Users report discomfort with the possibility that their voice assistants might capture private conversations, sensitive information shared within the home, or discussions that were never intended for the device [14][10]. Beyond concerns about unauthorized data collection, users express uncertainty about the broader data ecosystem surrounding voice assistants: which companies have access to their voice recordings, how long data is retained, what purposes it serves beyond responding to their immediate requests, and what control mechanisms exist to limit data collection or delete previously recorded information [7][10].

These privacy concerns are compounded by limitations in how users can access privacy information and exercise privacy controls for their voice assistants. While users primarily interact with voice assistants through speech for everyday tasks such as asking for weather updates, controlling lights, and playing music, they must switch to companion smartphone applications to review privacy policies or modify privacy settings. This modality mismatch creates a disconnect between the natural interaction paradigm and the mechanisms provided for privacy management. When users do navigate to these companion applications to review privacy information, they frequently encounter lengthy privacy policies filled with legal and technical jargon that obscure rather than clarify data practices. Our work also confirms that the complexity and inaccessibility of these text-heavy documents often lead users to feel overwhelmed and resigned, with some expressing sentiments such as "I think I anyway have to agree with it." This resignation suggests that current privacy notice and choice mechanisms fail to meaningfully empower users to understand and control their privacy, instead creating barriers that discourage engagement with privacy information altogether.

The combination of incomplete mental models regarding the device operation, complex privacy policies, and modality mismatches in privacy control access creates a significant research gap. Given that voice represents the primary interaction modality for voice assistants, and given that current privacy notice and choice mechanisms rely heavily on visual, text-based interfaces that users find difficult to navigate and understand, there is a clear opportunity to explore whether delivering privacy notices and enabling privacy choices directly through the voice channel could improve transparency and user control. Voice-based privacy interactions could potentially increase users' exposure to privacy information by delivering it through the same modality they already use for everyday assistant tasks, reduce barriers to accessing privacy controls by eliminating the need to switch to companion applications, and present privacy information in conversational formats that may be more accessible than dense legal text. However, prior research has not systematically examined user perspectives on voice-based privacy notices and choices for voice assistants, nor evaluated whether such mechanisms can effectively increase transparency and user confidence in data practices.

To address this gap, our research investigates three questions:

- 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 voice user interface for privacy notices increase the transparency of privacy notices and choices for users and, in turn, build more trust with voice assistants?

To answer these questions, we conducted a semi-structured interview study with voice assistant users. The study combined open-ended background questions about participants' current experiences, privacy concerns, and existing practices with a structured evaluation of a functional voice user interface prototype. The prototype, branded as "Nexa," demonstrated how privacy notices could be delivered and privacy choices could be exercised through voice interactions. To enable systematic comparison between voice-based and graphical approaches, participants evaluated the same privacy tasks using both a voice user interface and a graphical user interface in counterbalanced order. This design allowed us to identify not only what privacy concerns users face with current implementations, but also how voice-based privacy mechanisms compare to traditional graphical alternatives in terms of user preference, perceived transparency, and confidence in data practices.

This research contributes to the growing body of work on voice assistant privacy by documenting specific pain points users encounter when attempting to understand and control privacy with current implementations, eliciting detailed user perspectives on the benefits and limitations of voice-based privacy notices and choices, and providing empirical evidence about whether voice-based privacy transparency mechanisms can address gaps left by text-based policies and application-based controls.

# 3 Related Work

## 3.1 Existing user-privacy and trust perceptions about voice assistants

The proliferation of smart voice assistants has generated substantial research examining user privacy concerns and trust perceptions. A systematic review by Seymour et al.[13] identified privacy as the most prevalent ethical concern investigated in voice assistant research, with studies revealing that users express significant apprehension about data collection and processing by vendors and third parties. Users frequently distinguish between acceptable and unacceptable data use based on factors such as first-party versus third-party access, the sensitivity of subject matter, and whether voice recording data is analyzed by humans or machines.

A critical finding across the literature is that users possess incomplete and often inaccurate mental models of how voice assistants operate. Major et al. [9] demonstrated this through their survey of 237 Alexa users, finding that 62.4% of participants incorrectly believed that everything Alexa says is programmed by Amazon, unaware that third-party developers can build skills with direct access to user data. Surprisingly, users with greater familiarity and experience with Alexa were more likely to mistakenly assume third-party skills were native Amazon functionality, suggesting that prolonged use may foster a false sense of understanding, with users developing misplaced confidence in their knowledge of the system rather than genuine privacy awareness.

This lack of understanding extends to privacy controls. Seymour et al. [13] reported that users often exhibit poor knowledge of available privacy mechanisms, such as the ability to replay and delete stored voice recordings. Ahmad et al. [2] found that users frequently distrust software-based privacy controls, viewing them as potentially deceptive, and instead prefer "informal" coping strategies such as unplugging devices entirely to ensure their privacy. Their experimental study with 261 participants demonstrated that hardware-based tangible controls significantly increased perceived trust, reliability, and sense of control compared to software-based mechanisms.

The uncertainty surrounding device behavior, particularly regarding when recordings are stored and transmitted, compounds these trust deficits. Users express skepticism about manufacturers' honesty given their vested interests in data collection, with many adopting a resigned acceptance of privacy tradeoffs in exchange for device functionality and convenience.

## 3.2 Guidelines for Design and Implementation of VUIs

Murad et al. (2023) [11] synthesized 336 VUI design guidelines from academic literature into 14 core principles through a rigorous meta-analysis of 40 peer-reviewed papers spanning 28 conferences and journals from 1995 to 2021. The most emphasized guidelines focus on providing interactive experiences with substantial user control (158 guidelines) and designing natural conversations aligned with real-world dialogue patterns (130 guidelines). Clear, informative feedback (99 guidelines) and robust error handling (72 guidelines) also emerged as critical. Key recommendations include transparency about system intelligence and capabilities, breaking tasks into manageable sub-components to reduce cognitive load, personalizing experiences based on context, and maintaining consistency across interactions. Notably, accessibility and diversity received minimal attention (7 guidelines), representing a significant research gap.

Among the most prominent principles, the guideline on user control emphasizes making it easy for users to identify different areas of the interface, interrupt the VUI when desired, and navigate back to main menus at any point. The authors note that providing customizable commands and shortcuts can significantly enhance usability. The second most discussed principle—designing conversations that map to real-world norms—addresses the importance of using appropriate prosody, clear pronunciation, and low latency in system responses. This principle also advocates for abiding by natural turn-taking protocols, providing brief system responses, and allowing for user interruptions throughout the dialogue tree. The feedback guideline stresses providing immediate confirmation when the system processes user input, making system status visible throughout interaction, and delivering unambiguous feedback about recognition failures. For error handling, the synthesis recommends communicating system functionality and potential limitations upfront to prevent errors, providing clear error messages when failures occur, and offering mechanisms for conversation repair, such as undo functions and clarification prompts

### 3.3 Privacy and Transparency Guidelines for VUI Design

While general VUI design principles establish foundational interaction patterns, privacy and transparency have emerged as particularly critical considerations given the always-listening nature of voice assistants. Research demonstrates that users express significant concerns about such devices, yet often fail to engage with available privacy controls. Khezresmaeilzadeh et al.[5] found that voice assistant profiling can occur without user interaction and may be inaccurate, highlighting the need for transparent data practices and opt-out mechanisms. Liu et al.[8] showed that while users prefer stronger privacy protections when installing voice assistant apps, the majority do not review permissions in detail, suggesting that privacy information must be made more salient and accessible during onboarding and throughout usage.

Studies on proactive assistants reveal that users prioritize control over their data and the assistant's actions, particularly for sensitive information like financial details. Ahmad et al. [2] demonstrated that hardware-based tangible controls, such as physical mute switches with visible feedback, significantly increase perceived trustworthiness and usability compared to software-based mechanisms, which users view as potentially deceptive. These findings collectively suggest that effective voice assistant design should incorporate transparent data practices, accessible and meaningful privacy controls, and tangible feedback mechanisms that provide users with assured confidence about device state.

### 3.4 Large Language Models as an Enabling Technology for Enhanced VUIs

Despite the comprehensive design guidelines established in the literature, traditional voice assistants have struggled to implement many of these recommendations due to technological constraints, especially in maintaining context. Large Language Models have emerged as a transformative technology that can address many of these limitations. Klein et al. [6]note that conventional VUIs struggle with context-sensitive interaction and natural conversation flow, but ChatGPT and similar models can learn from human dialogues to enable more accurate, context-aware responses.

Recent implementations demonstrate LLM advantages across diverse domains: Wang et al. [15] developed an LLM-powered conversational system for the social robot Haru that generated contextually appropriate responses with expressive behaviors, receiving positive feedback for empathy and naturalness. Similarly, Huang et al. [4] found that LLM-based voice assistants like "Driver Mate" offered more intelligent natural conversations compared to predecessors like Siri or Alexa, with high user acceptance. Yang et al.'s Talk2Care system [17] leveraged GPT-3.5's multi-turn conversation capabilities for personalized healthcare communication, achieving good usability scores. However, challenges persist, including hallucinations, repetitive responses, and ethical violations, requiring careful prompt engineering and human oversight.

Building upon this body of literature, our work extends LLM capabilities to address the privacy transparency gap in current voice assistants. We develop a VUI prototype that incorporates the key design principles identified in the literature: leveraging the contextual understanding and natural conversation capabilities of LLMs to deliver privacy information in a manner aligned with real-world dialogue patterns, implementing a guided onboarding phase that proactively communicates data practices, and providing clear opt-out mechanisms accessible through natural voice commands. We acknowledge that, as a software-based prototype lacking hardware controls such as a physical mute button, our system cannot fully address user skepticism toward software mechanisms and that LLM hallucinations remain a potential concern. By synthesizing the general VUI guidelines emphasizing user control and transparency with the privacy-specific recommendations of improving the accessibility of privacy notices, our prototype aims to demonstrate how LLM-powered voice assistants can enhance the clarity of understanding privacy notices and the ease of executing privacy choices.

# 4 Competitive Analysis

To inform the design of our voice assistant prototype and establish a baseline understanding of current industry practices, we conducted a competitive analysis of three leading voice assistants: Amazon Alexa (Echo Pop), Google Home, and Apple HomePod Mini.

Our analysis focused specifically on evaluating each assistant's capability to respond to privacy-related questions through voice commands. We examined whether users could obtain information about data collection practices, access privacy settings, and manage their data using only voice interactions, without resorting to their companion apps or websites. By systematically testing similar privacy and queries across all three platforms, we sought to characterize current limitations and inform alternative design approaches that will guide the design principles of our VUI prototype.

The privacy questions assessed three core dimensions—information transparency (e.g., "What data do you collect from me?"), user control (e.g., "Can I change my privacy settings through voice?"), and data management (e.g., "Delete my voice recordings") while non-privacy questions served as control items to establish baseline functionality.

All three devices were configured with standard user accounts and default privacy settings. Testing was conducted in October 2025 in a controlled environment, with two researchers with different accents present during interactions to mitigate potential accent bias in voice recognition. All interactions were recorded via Zoom and transcribed. Detailed findings for each assistant are documented in Sections 4.1, 4.2, and 4.3, with a concise comparison of key responses presented in Table 1.

## 4.1 Alexa

We asked a total of 47 questions, comprising 32 privacy-related questions and 15 non-privacy-related questions. We observed that the device demonstrated moderate performance on privacy transparency, with a consistent pattern of redirecting users to external resources rather than providing voice-based controls. The device successfully performed deletion tasks through voice commands and answered certain informational privacy questions by reading from its privacy policy, but could not execute most privacy setting changes through voice. For non-privacy-related questions such as general knowledge ("Who is the president of China?"), weather, and time queries, the device performed reliably and provided accurate responses.

We began our interaction by asking questions about privacy settings. When prompted with "Walk me through my privacy settings," the device consistently responded by directing users to the Alexa app or Amazon.com/alexaprivacysettings and sent a link to the Alexa app. This same response pattern was triggered for variations including "How can I change my privacy settings?" and "Can I change my privacy settings through voice commands?" The device did not distinguish between requests for information about privacy settings versus requests to actually change settings through voice—all received the same redirect response. When explicitly asked, "Why can't I change my privacy settings through voice?", the device simply replied, "Hmm, I don't know that."

For data deletion, the device demonstrated notable capability with granular control. The command "Delete what I just said" successfully triggered the deletion of recordings from the last 10 minutes, with Alexa confirming "Okay, I'll delete any recordings from the last 10 minutes." The device also supported broader time-based deletion—asking "Delete all my voice recordings today" prompted a confirmation ("You'd like to delete the recordings of everything said to me today, is that right?") and upon confirmation, the deletion was executed. However, deletion commands sometimes failed inconsistently; attempts to delete recordings occasionally triggered an irrelevant response about connecting video devices ("To control a video device with your voice, you need to connect it first"), demonstrating unreliable command recognition for certain phrasings of deletion requests.

Regarding voice recording information, the device provided accurate responses when asked as knowledge questions. For "How can I delete my voice recordings?" it correctly directed users to Amazon.com/alexaprivacysettings or the privacy section of the Alexa app. When asked about what information is collected, the phrasing mattered significantly: "What information do you collect from me?" retrieved an unhelpful response from Quora.com, while "What kind of information do you collect from me?" triggered a proper response: "By default, I only send audio back to Amazon when I hear you say the wake word. For more information and to view Amazon's privacy notice, visit the Help section of your Alexa app."

A notable issue emerged with voice recording storage queries. When asked, "Where are my recordings saved?", the device searched the web and provided an irrelevant answer about Skype recordings from apeaksoft.com, completely misunderstanding the query. Similarly, "Where are my interactions with you saved?" received "I'm not quite sure how to help you with that," indicating the device could not provide information about its own data storage practices.

For interest-based advertising questions, the device demonstrated different behavior depending on how the question was framed. When asked, "Can I opt out of the internet-based ads recommended by you?" as a knowledge question, the device read the relevant section from Amazon's privacy policy: "Even if you choose not to see interest-based ads, you may still see personalized product recommendations and other similar features on Amazon..." However, when asked to actually perform the opt-out action ("Can you opt me out of all the advertising from my Amazon account?"), The device could not execute it, responding, "I'm not quite sure how to help you with that."

The device retrieved information from external web sources with mixed accuracy. For "Can I opt out from sending my recordings to Amazon?", the device fetched a response from Entrepreneur.com discussing changes to Amazon's voice recording policies, which was tangentially related but not directly actionable. When asked about turning off voice recognition, it retrieved helpful information from Storables.com explaining how to disable the "Use Voice Recordings" option in the Alexa app, but could not execute this change through voice.

The device performed well on certain functional tasks. It successfully demonstrated how to send an email by providing step-by-step instructions: "To link your email to your Alexa device, open the Alexa app, go to the Settings menu..." It also correctly answered "Are you listening right now?" with a comprehensive response: "I'm designed to protect your privacy. On Echo devices, you'll always know when Alexa is listening to your request and sending it to Amazon's secure cloud. Just look for the blue light indicator or listen for an audio tone." Additionally, the device could access Amazon order history, correctly listing the last three purchases when asked.

## 4.2 Google Home

We asked a total of 52 questions, comprising 38 privacy-related questions and 14 non-privacy-related questions. We observed that the device demonstrated mixed performance on privacy transparency: it answered some privacy-related questions satisfactorily (approximately 12 of 38), while others received redirects to external resources, irrelevant answers, or outright failures to understand. For non-privacy-related questions such as weather queries ("What's the weather in Pittsburgh?"), The device performed well and provided accurate, immediate responses. The device also successfully performed calendar-related tasks such as reading agenda items, but consistently denied access to emails and messages, stating "Sorry, I can't send messages yet" or "Sorry, I don't understand."

We began our interaction by asking questions about privacy settings. When prompted with "Walk me through my privacy settings," the device consistently redirected users to myaccount.google.com and sent a link to the connected phone. This behavior triggered reliably whenever the phrase "privacy settings" was detected, representing a standardized response pattern. However, when we asked, "Can I change my privacy settings through voice?" the device provided an irrelevant answer about adaptive brightness ("I can't change adaptive brightness on other devices yet"), demonstrating a failure to understand the actual query. Similarly, when asked about "privacy choices," the device redirected to the website rather than providing voice-based information.

For data deletion, the device demonstrated notable capability. When asked to "delete what I just said," it successfully removed recent assistant activity and confirmed the deletion. Time-based deletion also worked effectively—asking "delete my recordings for the past 10 minutes" resulted in the device confirming deletion of activity within that specific timeframe. However, for broader deletion requests such as "delete all my voice recordings," the device redirected users to the Google Home app's My Activity page rather than performing the action directly.

Regarding data collection transparency, when asked "What information do you collect from me?" the device provided information about Google's general data collection practices and redirected to safety.google, but notably failed to provide information specific to Google Home or the smart speaker itself. When we attempted to ask, "What kind of information do you collect from me?" or "What data do you, as a speaker, collect from me?", the device either did not understand or provided irrelevant responses about music playback.

The device also consistently failed to answer "How are my voice recordings being used?", responding with "Sorry, I don't understand" across multiple attempts.

A notable pattern emerged with voice recording queries. When asked, "Can I access my voice recordings through you?" or "Do you use my voice recordings?" the device misinterpreted these as requests for note-taking functionality, responding with "Sorry, I can't record your voice, but you can ask me to take a note for you." This semantic confusion prevented users from obtaining information about existing recordings. However, when asked "Are you storing my voice recordings?", the device provided a helpful redirect to g.co/assistant/yourdata with a link sent to the phone. For questions about opting out of voice recordings ("Can I opt out from sending my recordings to Google?"), The device simply responded, "Sorry, I don't understand."

The device performed well on certain instructional queries. When asked, "How can I opt-in voice match?", it provided clear step-by-step instructions for accessing Voice Match settings in the Google Home app. Similarly, "How can I opt out from personalized ads?" received a comprehensive response with specific instructions from policies.google.com. The device also correctly answered "Are you always listening to me?" with an informative response: "I'm designed to wait in standby until I detect an activation, like when I hear, hey Google. The status indicator on your device lets you know when I'm activated." Questions about Guest Mode in Google Home were also answered satisfactorily, with the device explaining that audio recordings and assistant activity would be deleted and no longer tied to anyone while in this mode.

We observed voice verification requirements for certain privacy-sensitive operations. When a different speaker attempted to delete voice recordings, the device requested voice verification, stating, "I can only share personal information once I verify your voice." However, upon repeated attempts with the same accent, the device eventually performed the deletion without completing verification, suggesting inconsistent enforcement of voice match security.

For queries about deleting "loaned data" (clips used to improve Google's services), the device requested permission to allow "assistant personal results" on the device, requiring users to change settings in the Google Home app before proceeding. When asked about stored voice profiles, the device provided an irrelevant response ("My team at Google gave me my voice, but you give me the best reason to use it") rather than information about user voice profiles. The device appropriately refused to disclose account passwords, stating "Sorry, I can't tell you your password" and redirecting to passwords.google.com.

Throughout the interaction, we observed that some commands were not well understood, regardless of whether they were privacy-related, suggesting general limitations in the device's natural language comprehension. The device demonstrated a clear pattern of providing "how-to" instructions effectively while struggling with queries requiring direct information disclosure or voice-based privacy controls.

### 4.3 Apple Homepod mini

We asked a total of 28 questions, comprising 19 privacy-related questions and 9 non-privacy-related questions. We observed that the device answered only one privacy-related question satisfactorily. The device either refused to answer or deflected 18 of the 19 privacy questions through various response patterns: direct refusals ("Sorry, I can't help you with that"), acknowledgments without answers ("Hmm, I don't have an answer for that"), redirects to the iPhone for web results, or generic website redirects. For non-privacy-related questions involving personal content such as messages and calendar access, the device repeatedly asked us to enable Find My on our iPhone or iPad, despite this feature already being enabled.

We began our interaction by asking five semantically similar questions about privacy settings (e.g., "Walk me through my privacy settings," "How can I change my privacy settings?"), all of which received the same refusal: "Sorry, I can't help you with that." We then asked questions about the types of information collected from the user, whether the device is actively listening, and data sharing practices, for which the standard response was "Hmm, I don't have an answer for that." Notably, the question "What kind of information do you collect from me?" received complete silence with no response at all. For questions pertaining to the deletion of conversation data, the assistant indicated it could show web results if we asked again from our iPhones, effectively shifting the burden to the user. The device also explicitly refused requests to turn voice recognition on or off, stating "I'm afraid I can't do that," demonstrating a complete absence of voice-based privacy controls.

The only satisfactory answer we received was regarding whether the device is recording, for which it

stated: "I respect your privacy, and only listen when you're talking to me. You can learn about Apple's approach to privacy on Apple.com." However, even this response redirects to Apple.com rather than a specific privacy policy link, reducing its actionability. Furthermore, a semantically similar follow-up question, "When are you listening?" received "I don't have an answer for that," revealing inconsistency in the device's ability to address related privacy queries.

Table 1: Summary of Voice Assistant Responses to Privacy and General Questions

| | **Question** | **Alexa** | **Google Home** | **Apple HomePod** |
|---|---|---|---|---|
| **Privacy Questions** | | | | |
| 1 | Walk me through my privacy settings | ✓Redirected to Alexa app or Amazon.com/alexaprivacy settings | ✓Redirected to myaccount.google.com | ✗"Sorry, I can't help you with that" |
| 2 | Delete my voice recordings | ✓Successfully deleted with confirmation | ✓"No problem. I deleted your most recent assistant activity" | ✗"I don't have an answer for that" |
| 3 | Are you listening right now? | ✓Explained privacy design and blue light indicator | ✓"You have my full attention" / Explained standby mode | ✓Answered correctly with "I respect your privacy, and only listen when you're talking to me. You can learn about Apple's approach to privacy on Apple.com" |
| 4 | Where are my recordings stored? | ✗"I'm not quite sure how to help you with that" | ✗"Sorry, I don't understand" | ✗"Sorry, I couldn't find that in your home" |
| 5 | Can I opt out of internet-based ads? | ✓Read relevant section from the privacy policy | ✓Redirected to "policies.google.com" and provided further steps | – Not tested |
| 6 | Can I change my privacy settings through voice? | ✗Redirected to app (no voice control) | ✗Misunderstood the question | ✗"Sorry, I can't help you with that" |
| 7 | What data do you collect from me? | ✓Answered it by clearly mentioning that it only sends audio back to Amazon when it hears the wake word. | ✗Gave a correct answer for the Google account in general, but when asked about the speaker, it responded with "Sorry, I didn't understand" | ✗Silent/no answer |
| 8 | Do you use my voice recordings? | ✓clearly mentions about the human review of voice recordings | ✗Completely misunderstood | ✗"I don't have an answer for that" |
| 9 | Can I opt out from sending my recordings? | ✗Redirected to app (no direct opt-out) | ✗"Sorry, I don't understand" | ✗"A visit to Apple.com should help" |
| **Non-Privacy Questions (Control)** | | | | |
| 10 | What's the weather/temperature? | ✓Provided current temperature and forecast | ✓Provided detailed weather information | – Not tested |
| 11 | Access my calendar | ✓Successfully retrieved calendar events | ✓Successfully retrieved calendar events | ✗Failed (required Find My setup) |

## 4.4 Takeways

Our competitive analysis revealed systematic exclusion of privacy management from voice interfaces across all three platforms. Only recording deletion was achievable via voice, and only on Alexa and Google Home. Privacy functions such as understanding data collection, accessing settings, and opting out of data uses required navigating companion apps or websites. Privacy questions, such as where recordings are stored, what data is collected, how recordings are used, whether users can opt out of data collection or advertising, and whether privacy settings can be changed via voice, were not answerable by all the VAs. While, Alexa

redirected users to external resources, Google misunderstood or provided irrelevant responses, and Apple refused engagement entirely. For our VUI prototype, these findings underscore the need for voice-first privacy management where all privacy functions are fully accessible through voice commands, with precise responses and the ability to maintain multi-turn conversations.

# 5 User Study Methods

## 5.1 Recruitment

Prior to formal participant recruitment, we conducted three rounds of informal pilot surveys that help us finalize the interview script. Participants in these pilots were members of the research team, friends, and colleagues recruited via personal connections. We used feedback from each round to iteratively revise the survey's wording, structure, and clarity. No compensation was provided to participants in these informal pilots.

For formal recruitment, we employed multiple channels to reach potential participants. Recruitment materials included physical flyers distributed in public spaces and personal outreach to friends and colleagues. All recruitment materials directed interested individuals to complete a screening survey to assess their eligibility for participation.

The screening process was implemented through a multi-stage approach using the Qualtrics online survey platform. Potential participants first encountered the screening survey, which consisted of seven questions designed to assess eligibility criteria and gather preliminary information about voice assistant usage patterns. The screening survey assessed age, geographic location, English language proficiency, prior experience with voice assistants, and recent usage frequency. Specifically, participants were asked to report how many times they had used a voice assistant in the past three months, with response options ranging from zero times to 15 or more times. The final questions in the screening survey asked participants about their willingness to participate in the interview study and requested their email addresses for follow-up contact.

Participants were eligible for the study if they met the following criteria: adults aged 18 years or older, currently located in the United States, able to read and speak English, prior experience using voice assistants, including Amazon Alexa, Google Home, Apple Siri, or similar products, and a minimum usage threshold of at least 10 times in the past three months. These eligibility requirements were designed to ensure participants had sufficient familiarity with voice assistant technology to provide meaningful insights during the interview.

Eligible participants were contacted via email with an invitation to participate in the interview study. The email invitation outlined a two-step enrollment process. First, participants were directed to complete an online consent form detailing the study procedures, risks, benefits, and data handling practices. Second, after providing consent, participants were provided with a Calendly scheduling link to book their interview session at a convenient time. This automated scheduling system allowed participants flexibility in selecting interview slots while streamlining the coordination process for the research team.

Throughout the recruitment process, participants were informed that their participation was entirely voluntary and that they could withdraw at any time without the fear of loosing the compensation. Contact information for the principal investigator and the institutional review board was provided in all recruitment materials and consent documentation, allowing participants to ask questions or raise concerns at any stage of the study.

## 5.2 Research Ethics

This study received approval from the Carnegie Mellon University Institutional Review Board (Study ID: STUDY2025_00000334). All interviews were conducted remotely via Zoom to accommodate participants across the United States. Participants were not required to enable their cameras, minimizing potential discomfort with video recording while still allowing for natural conversation.

Informed consent was obtained from all participants prior to the interview. Participants reviewed and completed an online consent form that detailed the study purpose, procedures, risks, benefits, data handling practices, and their rights as research participants. At the beginning of each interview session, researchers verbally summarized key consent information, including the 60-minute duration, voluntary nature of participation, audio recording and transcription procedures, and compensation details. Participants verbally confirmed their consent to participate and to be recorded before the interview commenced. Participants were informed of their right to withdraw from the study at any time without penalty and without forfeiting their compensation.

All interview sessions were audio-recorded and transcribed using Zoom's recording and transcription features. Participants were instructed to avoid sharing personally identifiable information during interviews. Audio recordings and transcripts were stored on password-protected, access-restricted cloud storage, with

access limited to members of the research team. All identifying information was removed from transcripts and research outputs. Only pseudonyms or participant identifiers are used in research reports and publications. Participants received a $30 Amazon gift card following completion of their interview session.

## 5.3 Interview Design

The interview study employed a semi-structured interview format combining open-ended background questions with structured prototype evaluation tasks. Each interview session lasted approximately 60 minutes and was conducted remotely via the Zoom video conferencing platform. The interview design was structured to address three primary research questions: understanding privacy concerns with current voice assistant privacy notices and choices, exploring user perspectives and preferences for voice-based privacy interactions, and evaluating whether a designed voice user interface for privacy notices could increase transparency and trust. The full interview script is available at Appendix **??**

### 5.3.1 Interview Structure and Components

The interview protocol consisted of two major components conducted sequentially within each session. The first component focused on participants' background experiences, current practices, and attitudes toward voice assistants and privacy. The second component involved the evaluation of a functional prototype system demonstrating enhanced privacy notice and choice mechanisms through both the VUI and the GUI. This dual-component structure allowed researchers to first establish a baseline understanding of participants' existing experiences and concerns before introducing novel interface designs for comparative evaluation.

To control for potential order effects in the prototype evaluation component, participants were randomly assigned to one of two counterbalanced conditions. Half of the participants interacted with the graphical user interface first, followed by the voice user interface. The remaining half experienced the reverse order, beginning with the voice user interface and then proceeding to the graphical interface. This counterbalancing approach ensured that any observed differences in user responses or preferences could be attributed to interface characteristics rather than learning effects or fatigue from repeated task completion.

### 5.3.2 Background and Experience Questions

The initial portion of the interview gathered comprehensive information about participants' prior experiences with voice assistant technology. This section began with broad questions about voice assistant usage patterns, including which voice assistant platforms participants had used, the types of devices they accessed these assistants through, and the contexts in which they employed voice assistant functionality in their daily lives. Questions explored both the situations where participants actively chose to use voice assistants and circumstances where they deliberately avoided using these technologies. These questions directly addressed the first research question by uncovering existing pain points and concerns in participants' everyday interactions with voice assistants.

Recognizing that voice assistants are frequently deployed in shared household environments, the interview protocol included specific questions about shared device usage. Participants were asked whether they used voice assistants that were accessible to multiple household members, how many people in their household used the shared device, and whether they experienced any privacy considerations when using shared voice assistants. These questions elicited information about contextual privacy concerns that might not arise with personal devices, contributing to an understanding of the full spectrum of privacy issues users encounter with current voice assistant implementations.

The background section also explored participants' interaction style preferences with voice assistants. Participants were asked whether they preferred to give short, directive commands or longer, more conversational queries when interacting with voice assistants. Follow-up questions probed whether these preferences varied depending on the type of task being performed or the environment in which the voice assistant was being used. Additionally, for participants who had experience with AI chatbot systems that offer voice input and output capabilities, questions explored perceived differences between interacting with AI assistants versus traditional voice assistants. These questions established a baseline understanding of participants' communication preferences and mental models, which informed the interpretation of their responses to the prototype voice interface designs.

### 5.3.3 Privacy Awareness and Current Practices

Following the general background questions, the interview transitioned to focus specifically on privacy-related awareness and behaviors. This section was organized into four distinct subsections, each addressing different aspects of how users currently engage with voice assistant privacy mechanisms.

The first subsection examined participants' engagement with privacy policies for voice assistant platforms. Participants were asked whether they had ever read privacy policies for their voice assistants, and if so, which platforms' policies they had reviewed and in what scenarios they chose to read these documents. Questions explored participants' motivations for reading privacy policies, whether any aspects of the policies surprised them, and their overall user experience in finding and reading privacy information. For participants who reported positive experiences, questions probed which design elements or features made the process easier or helped them feel more in control of their data. Conversely, for participants who found the process difficult or chose not to engage with privacy policies at all, questions explored the barriers and reasoning behind these choices. These questions directly informed the first and second research questions by revealing current pain points with text-based privacy policy presentations and establishing baseline preferences that could inform voice-based privacy notice design.

The second subsection focused on privacy settings management. Participants were asked about their awareness of available privacy settings on their voice assistants and whether they had ever modified these settings. For participants who had changed privacy settings, questions explored which specific settings they modified, their reasons for making changes, their opinions about default privacy configurations, and their experiences navigating the settings interface. Follow-up questions asked participants to identify features or design elements that made the process easier or more difficult, and whether any aspects of the experience increased or decreased their sense of control over their data and certainty about information collection practices. For participants who had not changed privacy settings, questions explored their reasons for maintaining default configurations. Additionally, participants were asked whether they had attempted to change privacy settings through voice commands rather than graphical interfaces, what commands they had tried, whether they succeeded, and their assessment of the ease or difficulty of voice-based settings modification.

The third subsection addressed data access and control practices. Participants were asked whether they had ever checked to see what information their voice assistant had collected from them. For participants who had accessed their data, questions explored which types of collected data surprised them, the scenarios that motivated them to check their data, and their user experience in locating and reviewing the collected information. Similar to the privacy settings questions, follow-up probes identified design elements that increased or decreased the ease of the process and affected participants' sense of certainty about information collection practices. Questions also explored whether participants had attempted to delete their conversation data, the methods they used to perform deletion, and whether they were aware that data deletion could be performed through voice commands. For participants who had not accessed their collected data, questions explored the reasons for this choice.

The fourth subsection examined voice-based privacy interactions with current voice assistant systems. Participants were asked whether they had ever tried to ask their voice assistants privacy-related questions, such as inquiries about data collection practices or requests to change privacy settings. For participants with this experience, questions explored what specific questions they asked, how the voice assistant responded, their motivations for using the voice channel, the scenarios in which they made these requests, and their overall assessment of the experience. For participants without experience making voice-based privacy requests, questions asked them to imagine how such an experience might feel and whether they would be interested in asking similar questions if the capability were available in their daily life. All participants were asked about scenarios in which they might want to make voice-based privacy requests. These questions directly addressed the second research question by establishing current practices, unmet needs, and user expectations for voice-based privacy interactions.

### 5.3.4 Prototype Evaluation

The second major component of the interview involved evaluation of a functional prototype system called "Nexa," designed to demonstrate enhanced privacy notice and choice mechanisms for voice assistants. The prototype was specifically developed to test whether providing privacy information and controls through

voice-based interactions could address the limitations and friction points identified in current voice assistant privacy implementations.

The prototype evaluation centered on three standardized privacy tasks that represented common privacy-related information needs and control actions: understanding what types of personal data the voice assistant collects from users, opting out of interest-based advertising, and deleting stored voice conversation data. These three tasks were selected because they correspond to fundamental privacy rights related to transparency, choice, and control.

For each of the three privacy tasks, the prototype system offered two interface modalities: a graphical user interface implemented as an interactive Figma prototype, and a voice user interface implemented as a functional voice interaction system. The graphical interface represented an enhanced version of current smartphone application-based privacy settings, which was inspired by Amazon's Alexa settings page. The voice interface represented a novel interaction paradigm where users could ask natural language questions about privacy and receive spoken explanations, or issue voice commands to modify privacy settings with conversational confirmation dialogs.

During the graphical interface evaluation, participants were asked to use the think-aloud protocol while interacting with the Figma prototype. Participants were instructed to verbalize their thoughts continuously as they navigated the interface and completed tasks, including stating what they were looking at, what they were trying to accomplish, what they expected would happen, and anything that confused them.

For each privacy task in the graphical interface, participants were asked a series of structured follow-up questions. These questions asked whether anything made the process easy to understand or find, whether anything made the process difficult, and whether, for that particular topic they would prefer to check the information in the application or have the voice assistant tell them through voice. These questions provided task-specific usability feedback and began to establish comparative preferences between modalities.

During the voice interface evaluation, participants were presented with the three privacy task topics and asked to phrase natural voice commands or questions as they normally would when speaking to a voice assistant. The prototype system was implemented to continuously listen for participant speech and provide spoken responses. To avoid confusion during the voice interaction, participants were instructed that researchers would not be able to answer procedural questions through voice during the interaction, but could respond to questions sent via Zoom chat. This ensured that the prototype captured naturalistic voice interaction patterns without interference from researcher's speech triggering the system.

For each privacy task in the voice interface condition, participants were asked structured follow-up questions parallel to those asked in the graphical condition. Questions explored whether the voice assistant's response met participants' expectations, whether the response provided all the information they needed, and whether, for that particular topic they would prefer to listen to the voice assistant's reply or check the information in the application. These questions allowed direct comparison of user satisfaction and completeness perceptions across the two interface modalities.

Following completion of all three tasks in both interface conditions, participants were asked a series of comparative questions that required them to reflect on their experiences across both modalities. These questions asked participants to identify their overall preference between speaking to the voice assistant versus using the application for privacy-related tasks, to explain what factors influenced their preference, and to consider whether their preferences might vary for different types of tasks or in different situational contexts. Participants were also asked which interaction method helped them feel more confident or less uncertain about how their information is collected and shared, and which method made them feel more in control of their privacy choices. These comparative questions directly addressed the second and third research questions by revealing user preferences and assessing whether the voice interface design successfully increased transparency, certainty, and perceived control compared to graphical alternatives.

The prototype evaluation concluded with open-ended questions asking participants to suggest improvements to the voice interaction experience, including aspects such as response content, response length, speaking style, and interaction flow. Participants were also asked whether they would want to use similar voice-based privacy capabilities in their daily lives if such features were available, in what scenarios they would use these features, and any other ideas they had for improving privacy notices and choices in voice assistants. These questions generated design recommendations and identified remaining gaps or concerns not addressed by the prototype implementation.

#### 5.3.5 General Interface Preference Questions

After completing the detailed prototype evaluation focusing on privacy tasks, the interview returned to broader questions about voice user interface versus graphical user interface preferences across different types of tasks. Participants were asked to reflect generally on what kinds of situations they believed voice interfaces work better than graphical interfaces, and conversely, what kinds of situations they believed graphical interfaces work better than voice interfaces. These general questions provided context for interpreting the privacy-specific preferences expressed earlier and helped identify whether any unique characteristics of privacy tasks influenced interface preferences differently from other types of tasks.

#### 5.3.6 Interview Conclusion

The interview protocol concluded with an open-ended question asking whether there was anything participants thought was important to consider regarding voice assistant privacy that had not been discussed during the session. This question provided an opportunity for participants to raise concerns, ideas, or experiences that might not have fit within the structured question framework but were nonetheless relevant to understanding the full scope of user privacy concerns and needs.

Following this closing question, researchers thanked participants for their time and contribution, stopped the audio recording, and confirmed the email address or mailing address where the participant preferred to receive their compensation gift card. This post-recording confirmation was to preserve participants' privacy by not capturing identifying information in the research recordings

#### 5.3.7 Connection to Research Questions

The interview design's two-component structure directly addressed all three research questions through complementary approaches. The background and current practices section addressed the first research question by systematically documenting privacy concerns across multiple dimensions of voice assistant use, including shared device contexts, privacy policy accessibility, settings management, data access, error experiences, and trust implications. This section establishes a baseline understanding of problems with current implementations.

Both the background section and prototype evaluation addressed the second research question by exploring user perspectives, preferences, and suggestions for voice-based privacy interactions. The background section captured current practices and unmet needs, while the prototype evaluation provided concrete examples of voice-based privacy mechanisms that participants could experience and critique, generating specific design feedback and preference data grounded in actual interaction experiences rather than hypothetical scenarios.

The prototype evaluation component addressed the third research question by implementing a functional voice-based privacy notice and choice system and comparing user experiences, satisfaction, perceived information completeness, and feelings of control and certainty between voice and graphical modalities. The comparative questions and the recurring focus on certainty and control throughout the privacy-related sections provided multiple measures relevant to assessing whether the voice interface design achieved its goals of increasing transparency and building trust.

### 5.4 Design of VUI Prototype

#### 5.4.1 Prototype Overview and Purpose

The voice user interface prototype, branded as "Nexa," was developed to demonstrate enhanced privacy transparency mechanisms through voice-based interactions. The prototype enabled participants to complete three core privacy tasks: understanding what types of personal data the voice assistant collects, opting out of interest-based advertising, and deleting stored voice conversation data. Beyond these structured tasks, Nexa also supported open-ended privacy questions, allowing participants to explore voice-based privacy information naturally.

### 5.4.2 Technical Architecture and Implementation

The Nexa prototype utilized Claude Haiku 4.5 as its backend AI model, deployed through an Android application on a researcher's device during interview sessions. This architecture mimicked the cloud-based processing model of commercial voice assistants. The AI backend was configured through detailed system instructions that defined Nexa's onboarding sequence, specified which privacy policy content to reference, and constrained responses to prevent hallucinations [Instruction prompt for Nexa can be found in Appendix A.1]. Critically, the instructions explicitly directed the model to rely only on the provided privacy policy and FAQ documents rather than external knowledge, ensuring all participants received consistent information.

As a simulation, Nexa did not actually modify real data or settings. When participants issued commands to delete recordings or opt out of advertising, Nexa verbally acknowledged completion without executing actual changes. This approach allowed evaluation of voice-based privacy controls without requiring a complete backend infrastructure.

### 5.4.3 Knowledge Base Development

The research team developed a knowledge base consisting of a privacy policy, terms and conditions, and an FAQ document, all systematically adapted from Amazon Alexa's publicly available materials. This grounded the prototype in real-world voice assistant privacy practices. The adaptation process removed Amazon-specific service integrations (e-commerce, music subscriptions, Prime services) while retaining content about voice assistant features, data collection, privacy controls, and third-party integrations. All references to "Alexa" were replaced with "Nexa." The resulting privacy policy covered voice recording practices, cloud processing, voice identification, third-party skills, data retention and deletion, and advertising opt-out mechanisms. The FAQ document addressed eight common privacy concerns, providing Nexa with structured responses to anticipated participant questions [The modified privacy policy and FAQs can be found in Appendix A.2].

### 5.4.4 Conversational Design and Interaction Flow

Each Nexa session began with a mandatory onboarding sequence. After greeting the user, Nexa immediately delivered a standardized privacy policy summary scripted word-for-word in the system instructions. This summary covered six key points: cloud-based voice recording and processing, user ability to review and delete recordings through the app, physical microphone disable buttons, third-party data sharing, manual review of recording samples for service improvement, and voice-based privacy control options. After the summary, Nexa asked if the user had questions.

The conversational flow distinguished between information requests and action commands. When participants asked questions about privacy practices, Nexa provided explanatory responses based on the privacy policy. When participants issued directive commands (e.g., "opt me out of ads" or "delete my recordings"), Nexa acknowledged completion with appropriate confirmations. For interest-based ad opt-outs, Nexa added the clarification: "You may still receive personalized recommendations and other similar features on Nexa. You may also receive ads delivered on Nexa, but they will just not be based on your interests." Throughout all interactions, Nexa supported follow-up questions while keeping responses concise to avoid overwhelming users.

### 5.4.5 Pilot Testing and Iterative Refinement

The research team conducted six pilot rounds with friends and colleagues before formal interviews. Pilot testing focused on four refinement dimensions: response length (ensuring conciseness without excessive detail), tone and speaking style (professional and warm), accent and natural speech patterns (clear pronunciation and cadence), and system instruction clarity (ensuring the AI model followed protocols consistently). Each pilot round resulted in modifications to either the system instructions or supporting documents. For example, early pilots revealed overly lengthy responses, prompting explicit brevity constraints in the instructions. This iterative process ensured the prototype delivered consistent, appropriate, and user-friendly privacy interactions for the formal study.

## 5.5 Design of GUI

We chose the Alexa application as our reference object for several reasons. First, since Alexa has a more than 60% market share in the US [1], the user experience with the Alexa app can, to some extent, represent the overall user experience when adjusting voice assistant-related settings. Second, the Alexa app has a relatively independent privacy policy and control interface. Amazon has a dedicated privacy policy and terms of service specifically for Alexa, as well as frequently asked questions about privacy settings for Alexa. Furthermore, all of Alexa's settings are controlled through the Alexa app, which allows us to fully understand the privacy notices and choices provided by Alexa and select a subset of these notifications and controls to design a similar application prototype for use in interviews. To simulate how participants perform the privacy tasks with an application with control applications, we designed a GUI application for Nexa. The GUI is designed based on Amazon Alexa's mobile application, which allows users to control their Amazon account settings, conversations with Alexa, and Alexa device settings.

To simplify the user experience, we designed the settings section of the Nexa app to simulate the settings design of the Alexa app. The starting point of the GUI user experience is a settings screen, which includes three different parts: a screen displaying some basic, non-interactive settings, including account settings and personal information; a dedicated Nexa Privacy page allowing users to opt out of interest-based advertisment services, improving Nexa service with their data, and delete their conversation data; and a screen containing the Nexa privacy policy, terms of service, and privacy FAQs. We designed the interface using Figma, allowing users to swipe between different screens, navigate, and toggle different settings.

When designing the interface and its elements, we tried to adopt a design approach similar to Alexa's, using the same naming and descriptions for settings, and the same location and depth for displaying these settings as in the Alexa application. In the initial design, we completely mimicked Alexa to construct our GUI prototype. However, after two pilot interviews, we found that participants spent far longer than expected searching for the privacy policy, sometimes even requiring hints from the researchers. Therefore, in the Nexa Privacy interface, we added a link to the privacy policy near the other privacy settings to make it easier for participants to find the privacy policy. The GUI screens can be visualized in Appendix A.5.

## 5.6 Data Analysis

We analyzed the interview data using an inductive analysis approach with emergent coding. We first developed an a priori codebook based on the interview protocol and our research questions to capture anticipated themes related to participants' experiences with voice assistants and privacy practices.

To support systematic analysis, all interview transcripts were segmented into consistent sections corresponding to the structure of the interview script. Each section was independently coded by two researchers using the initial codebook. During this process, researchers were encouraged to refine existing codes and introduce new codes when participants' responses did not fit the predefined categories, allowing themes to emerge from the data.

After completing double-coding for each section, three researchers met to discuss and resolve coding differences. Discrepancies were addressed through discussion until consensus was reached, and the codebook was iteratively updated to reflect shared interpretations. This collaborative process helped ensure consistency in coding and strengthened the reliability of the resulting themes. The final codebook can be visualized in Appendix A.4

# 6 Results

## 6.1 General User Experience

Before discussing participants' privacy-related experiences, we first describe their general use of voice assistants to provide context for the findings. The three most commonly used voice assistant platforms among our participants are Alexa, Google Home, and Siri. In the following, we outline participants' frequent tasks, whether their use of voice assistants is shared with others, whether they have modified any voice assistant settings, and whether they prefer long commands or short commands.

### 6.1.1 Primary tasks

Participants reported using voice assistants primarily for tasks that are simple, low-effort, such as setting alarms and timers, checking the weather, playing music, and setting reminders. Some participants also reported using the voice assistant for other tasks, including controlling other smart home devices, sending and reading text messages, making calls, and checking opening hours. Participants P3 and P8 particularly mentioned using the voice assistant to broadcast messages in the household. Participant P3 noted that:

> *'Hey G, can you broadcast a message that dinner's ready?' And it does it throughout all of our Google Homes, so the whole house hears that dinner's ready, and it kind of gives everybody the ability to come down to the kitchen. (P3)*

From participant P8:

> *Or, I will also broadcast a message to my son in his room to tell him that dinner's ready. (P8)*

They have multiple voice assistants in each family member's room. Therefore, they can broadcast messages through one voice assistant device to all other voice assistant devices in different rooms.

### 6.1.2 Avoid certain tasks for voice assistant

Generally, participants reported avoiding the use of voice assistants for tasks that involved detailed, complex, or in-depth questions. Participants expressed concerns that voice assistants might not accurately understand such questions and, as a result, could provide low-quality or incomplete answers. Beyond concerns about question complexity, some participants also described deliberately avoiding specific types of tasks when using voice assistants. For example, Participant P4 reported that she did not want the voice assistant to retrieve information from other applications, citing a lack of trust in the assistant's behavior.

Together, these accounts illustrate participants' selective avoidance of voice assistants for tasks they perceived as complex, important, or involving interactions beyond simple information retrieval.

### 6.1.3 Shared voice assistants

Sharing practices of voice assistant devices varied across participants. Several participants (P3, P4, P7, and P8) reported using voice assistants in shared household environments with family members or roommates, whereas others described their devices as primarily personal. When asked whether there were tasks they preferred to perform independently or privately on shared voice assistants, participant P3 indicated a desire for greater privacy. She mentioned that she sometimes wanted to conduct searches or make purchases using the voice assistant without other household members being aware of these activities.

### 6.1.4 Change voice assistant settings

Only a few participants reported having changed the default settings of their voice assistants. Among these participants, changes varied in purpose and scope. Participant P2 mentioned changing the assistant's voice from a female voice to a male voice. Participant P6 described adjusting settings to enable parental controls for a voice assistant used by her children, as well as configuring guest settings when visitors were present in her home. Participant P7 reported updating the voice match feature multiple times to improve the assistant's

ability to recognize his voice. He also described managing settings for Samsung Bixby, a voice assistant embedded in his smartphone, to allow it to operate even when the screen was locked.

In contrast, several participants reported limited awareness of voice assistant settings, noting that they did not know where or how to modify these settings.

#### 6.1.5 Voice command length preference

Most participants reported preferring to use short and concise commands when interacting with voice assistants. In contrast, only one participant (P4) described a preference for using longer and more conversational commands.

> *"Hello, can you please help me play this song?" Or, "Hello, I feel like listening to this..." Yeah, I usually don't use very concise or direct commands.(P4)*

Participant P4 associated this preference with how she perceived the role of the voice assistant. She described the assistant as a friend rather than a tool and mentioned occasionally asking playful or "silly" questions for enjoyment. In this sense, using longer commands reflected her tendency to anthropomorphize the voice assistant and engage with it in a more conversational manner.

By contrast, participants who preferred short commands cited several practical reasons. Some participants noted that the response timing of voice assistants made longer utterances inconvenient, as the system often responded before they had finished speaking (e.g., P1). Others reported that voice assistants tended to perform poorly when given longer or more complex commands, leading them to simplify their interactions (P2, P3, P7, P9). Participant P8 explicitly described viewing the voice assistant as a tool rather than a conversational partner, which motivated her to issue brief, task-oriented commands. Similarly, Participant P1 explained that using short commands helped ensure accurate speech recognition.

> *So I think the thing is that I have to speak in English and pronounce the words in a certain way for Alexa to hear the entire thing, so I prefer very short commands. (P1)*

### 6.2 Current Privacy Practices

Participants generally reported limited engagement with privacy-related tasks associated with voice assistants. Most participants indicated that they had little experience reading privacy policies, modifying privacy settings, accessing collected data, or deleting stored conversation records. Even among participants who had previously performed some of these tasks, many were unable to clearly recall their experiences or describe the specific difficulties they faced.

In this section, we describe participants' current privacy practices related to voice assistants and summarize the factors that shaped these practices.

#### 6.2.1 Privacy notices

Only two participants (P6 and P8) reported having read the privacy policies for their voice assistants. Participant P6 indicated that she had read the privacy policies for both her Google Home and Alexa devices, while Participant P8 reported reading the privacy policy for her Google Home device.

Participant P6 described reading privacy policies at multiple points, including during device onboarding, when privacy policies were updated, and when she had specific privacy-related questions. She explained that her motivation for reading privacy policies stemmed from concerns about cybersecurity risks and a desire to protect her financial accounts and personal information. Additionally, Participant P6 noted that she was raised in a family that adopted a cautious attitude toward technology, which influenced her approach to privacy-related practices:

> *my father was in the Korean War, and was active in the military during the Cold War, and I think his paranoia influenced me a little bit about, you know, what the government could do and things like that. And so I think some of it's just a personal experience.[...]And you shouldn't... you shouldn't just accept. You should always read the terms and conditions. (P6)*

For participants who had not read privacy policies, several reasons were reported. Some participants described privacy policies as too long and difficult to read, often containing legal jargon (P3, P5, P9). Others noted that they would agree to the privacy policy regardless in order to use the voice assistant, which reduced their motivation to read it in detail (P7, P11). Additionally, participant P4 reported that she had never been explicitly prompted to read the privacy policy (P4).

Participant P4 further explained that she assumed applications with similar functionalities would share equivalent privacy practices. As a result, she did not perceive a need to read the privacy policies for each individual application.

> *I'm a big voice note user in other tools, like WhatsApp, so I guess that's probably why it was, like, so natural for me starting using voice notes on this new tool, as well. And since I was already... confident with it. I was already familiar with it. It never raised, like, any type of alarm. (P4)*

#### 6.2.2 Privacy choices

A small number of participants (P3, P7, and P9) reported taking steps to limit data collection during the onboarding phase of their voice assistants, such as opting out of analytics-related settings. Among all participants, only Participant P6 described regularly deleting conversation data collected by the voice assistant, including removing conversation records after private interactions.

In contrast, several participants reported limited awareness of available privacy controls. Some participants indicated that they did not know where to modify privacy settings or access collected data, while others were unaware that such options, such as adjusting data collection settings or deleting conversation histories—were available at all. Participant P11 further noted that she did not perceive her interactions with the voice assistant as involving sensitive information and, as a result, expressed little concern about data collection. Participant P8 described feeling frustrated about the data collection practices of the company behind the voice assistant. She believed that these companies already possess extensive information about users, which led her to view actions such as accessing or deleting conversation data as ineffective.

### 6.3 User Experience - VUI

After completing privacy tasks with the Nexa prototype, participants shared their experiences with using voice-based interactions for privacy notices and choices. In this section, we present participants' overall satisfaction with the VUI, their perceptions of information adequacy, their willingness to use voice-based privacy mechanisms in real voice assistants, and their suggestions for improvement. We then discuss participants' experiences with the GUI before comparing modality preferences.

#### 6.3.1 Overall Satisfaction with VUI

The majority of participants (9 out of 11) reported satisfaction with the VUI for privacy tasks. Participants cited several reasons for their positive experiences. Multiple participants appreciated the concise nature of Nexa's responses (P5), noting that the system provided information without overwhelming them with excessive detail. Participants also valued the confirmation messages Nexa provided after executing privacy choices (P5, P10), which reassured them that their requests had been understood and acted upon. Some participants highlighted Nexa's effective voice recognition and understanding of their questions (P9), while others noted that the system provided thorough answers (P8) and supported follow-up questions (P2).

Participant P3 specifically mentioned that Nexa's responses were more detailed than those provided by her current Google Home device, suggesting that the prototype demonstrated improvements over existing commercial implementations. Participant P5 further explained that she appreciated not having to ask very specific questions to obtain the information she needed, as Nexa provided guidance during interactions.

However, not all participants were equally satisfied. Participant P1 expressed dissatisfaction with the VUI, noting that the system paused between sentences in a way that felt unnatural. She suggested that Nexa should phrase its sentences better to improve the conversational flow. Additionally, P1 felt that the VUI did not give her an option to request more detailed explanations when she wanted additional information.

Participant P2 offered a more nuanced perspective, indicating general satisfaction but noting that the onboarding phase felt incomplete. She explained that while Nexa was reactive in responding to privacy questions and answers, she would have preferred the system to be more proactive by asking guiding questions to ensure users received complete information about privacy practices.

#### 6.3.2 Information Adequacy

When asked whether they received adequate information through the VUI to understand privacy practices and make informed decisions, most participants (7 out of 11) indicated that the information provided was adequate. Participants offered several explanations for this perception. Participant P3 noted that Nexa's responses were more detailed compared to her experiences with Google Home. Participant P6 appreciated that Nexa acknowledged limitations in its responses while still providing clear information, which she found trustworthy. Participant P9 highlighted that the system offered clarifications and gave specific answers to her questions, which helped her understand privacy practices.

Participant P2 explained that while the prototype was reactive rather than proactive in delivering privacy information, this reactive approach still provided adequate information because it allowed her to ask follow-up questions based on her specific concerns.

However, several participants expressed concerns about information adequacy. Participant P1 felt that the VUI did not provide adequate information because it did not offer her an option to explore topics in greater detail when she wanted more comprehensive explanations. Participant P8, while generally satisfied with the VUI's thoroughness, noted that responses were not thorough enough regarding third-party data usage.

#### 6.3.3 Willingness to Use VUI for Privacy Tasks

A strong majority of participants (9 out of 11) expressed willingness to use voice-based privacy interactions if such features were available in their current voice assistants. Participants described various motivations for this willingness. Participant P6 noted that privacy policies change very frequently, and having the ability to ask voice-based questions would help her stay informed about current privacy practices without needing to re-read lengthy policy documents. Participant P7 appreciated the ability to ask specific questions through voice, which he found more efficient than navigating through graphical menus. Participant P9 explained that Nexa's responses made her more privacy-aware, suggesting that voice-based interactions could increase user engagement with privacy information. Participant P10 valued what she described as Nexa's pro-privacy interaction style, which encouraged her to explore privacy controls. Participant P11 indicated that the ease of changing privacy choices through voice increased her willingness to actually modify her privacy settings.

Participant P4 captured a broader sentiment about trust and usability

> *Yeah, I guess the main thing is trust at the end of the day.*

This quote reflects how voice-based privacy interactions might contribute to building user confidence in voice assistant data practices.

Participants also described specific scenarios in which they would use voice-based privacy features. Participant P9 mentioned she would use it during the initial days of setting up a new voice assistant and for executing privacy choices. Participant P4 indicated she would ask privacy questions to learn more about third-party data usage. Participant P6 expressed interest in asking about parental controls and guest mode settings. Participant P7 noted he would use voice-based privacy features after developing some familiarity with the system.

Despite the overall high willingness, some participants expressed hesitation or limitations. Participant P5 indicated she would not be willing to use VUI for privacy tasks, citing her preference to perform these tasks manually through the GUI. Participant P10, while generally willing, noted that she would hesitate to use voice for privacy interactions when speaking about sensitive information, suggesting privacy concerns about discussing certain topics aloud.

Participant P4 articulated a tension between ease of use.

> *it's just that it's, like, extra effort that nobody's really willing to do, even though it takes, like, seconds to do it, just because there's no clear or immediate benefit from it in... I mean, apparently, no?*

This quote highlights a broader challenge in privacy management: even when tools are relatively easy to use, users may not engage with them if the benefits are not immediately apparent.

#### 6.3.4 Suggestions for VUI Improvement

Participants offered limited explicit suggestions for improving the VUI design. As noted earlier, Participant P1 suggested that the system should phrase its sentences better to reduce awkward pausing between sentences. Participant P2's feedback about making the onboarding phase more proactive by asking guiding questions also represents a design suggestion for future implementations.

### 6.4 Overall Satisfaction with GUI

Participants also completed privacy tasks using a GUI designed for the Nexa app. This section describes participants' positive experiences with the GUI as well as navigation and usability challenges they encountered.

#### 6.4.1 Positive GUI Experiences

Several participants reported positive experiences with the GUI for privacy tasks. Participant P2 appreciated that the content in the privacy tab was well-organized, making it easier to locate relevant information. Participant P6 particularly liked the clean labels and the fact that the privacy policy had its own dedicated section, noting that this design made privacy information easy to find. Participants P7 and P11 both described the GUI as straightforward and clean, suggesting that the interface design successfully reduced visual complexity.

When completing specific privacy tasks, some participants noted positive experiences. Participant P1 reported good experiences both with opting out of interest-based advertising and with deleting voice recordings through the GUI. Participant P10 specifically appreciated that the privacy information was presented in a separate privacy tab rather than buried within general settings.

#### 6.4.2 Navigation and Usability Challenges

Despite positive experiences, some participants encountered navigation difficulties and usability challenges with the GUI. Participants P5 and P9 both expected to find privacy settings under a "Settings" tab rather than a dedicated "Privacy" tab. This mismatch between user expectations and actual interface organization caused initial confusion about where to locate privacy controls.

Participant P5 experienced additional confusion, expecting the privacy tab to contain both privacy notices and privacy choices. When she initially explored the interface, she found only the privacy policy and did not immediately recognize where to access privacy control options.

Participant P8 described age-related challenges with exploring the GUI. She explained that as an older user, she did not have enough energy or motivation to systematically explore all settings and menus. As a result, she tended to click randomly through the interface rather than following a logical exploration pattern, which made finding specific privacy controls more difficult.

### 6.5 Modality Preferences

After interacting with both the VUI and GUI for privacy tasks, participants shared nuanced perspectives about the relative strengths and limitations of each modality. Rather than expressing uniform preferences for one interface over the other, participants described how different modalities suited different tasks, contexts, and usage stages. In this section, we present participants' overall modality preferences, task-specific preferences for privacy notices and privacy choices, the reasons underlying these preferences, trust considerations, context-dependent usage patterns, hybrid approaches that combine both modalities, and design suggestions for future implementations.

### 6.5.1 Overall Modality Preferences

When asked about their general preference between VUI and GUI for privacy-related interactions, participants demonstrated considerable variation in their responses. Two participants (P4, P9) expressed an overall preference for VUI, while three participants (P1, P7, P8) indicated a general preference for GUI. However, these simple categorizations obscure the complexity of participants' actual preferences, as many emphasized that their choice depended on specific tasks, contexts, or stages of interaction.

Participant P6 articulated this context-dependence particularly clearly, explaining that the appropriate modality depends on the user's environment and circumstances:

> *Again, I think it depends on the situation. If I'm driving in my own vehicle, it would be a good safety feature to have it talk. There's other times that we're in a setting that requires quiet, a library, a house of worship, an office. Those are times that it might be better to read it quietly, so you're not distracting people around you. An example would be if you're a student, and you're in the library, and people are trying to study for finals. They don't want to hear Alexa blast in there trying to study a new quiet, so, I just think it depends on your environment. Are you in a private place? Your home or car? Are you in a public place? I think we want to be respectful about people around us in certain private situations. If I'm in a hospital room sick, and that actually happened the other day, I went to urgent care. The other patients don't want to hear Alexa making a bunch of noise, because the doctors need quiet for the treatment. So I think it depends on where you are in your environment and makes the move. (P6)*

This perspective suggests that rather than one modality being universally superior, the choice between VUI and GUI should adapt to the user's physical and social context, including considerations of safety (driving), social norms (quiet spaces), and privacy of the surrounding environment.

### 6.5.2 Task-Specific Preferences for Privacy Notices

When specifically considering privacy notices—the information about data collection, processing, and sharing practices—participants demonstrated clearer patterns of preference. Five participants (P1, P2, P3, P7, P8) indicated a preference for receiving privacy notices through the GUI, while four participant (P4, P9, P10, P11) preferred VUI for privacy notices.

Participants who preferred GUI for privacy notices cited several interconnected reasons. Multiple participants noted that the GUI provided more detailed information and enabled in-depth reading (P1, P3, P7). Participant P3 specifically described the GUI as "more detail oriented" compared to the summary provided by VUI. Participant P5 explained that she preferred reading over listening for privacy information, while Participant P2 elaborated on the limitations of voice for information-dense content:

> *But the thing is, I'm scared if it's a voice assistant feature. It would be too much information to process verbally, and I like to read it instead of hearing it. (P2)*

This observation suggests that the visual modality may be better suited for processing complex privacy information that users need to comprehend, compare, and retain.

In contrast, Participant P10, who preferred VUI for privacy notices, cited that the voice interface provided adequate information and was comprehensive. She appreciated that she did not need to find information herself through navigation, as the VUI proactively provided relevant privacy details in response to her questions.

### 6.5.3 Task-Specific Preferences for Privacy Choices

For privacy choices, the actions users take to modify privacy settings such as opting out of advertising or deleting recordings—participants demonstrated more varied preferences that often depended on the specific action being performed.

For opting out of interest-based advertising, Participant P3 preferred the GUI, reasoning that the app interface might offer more granular opt-out options. She also noted that being able to see the opt-out action

in the app provided a greater perceived sense of control. Conversely, Participant P10 preferred VUI for opting out, citing that it provided adequate information for this task.

For deleting voice recordings, Participants P3 and P7 preferred VUI, with P7 noting that it was easier to perform deletion through voice. However, Participant P11 preferred GUI for deletion because doing deletion in GUI offered her more confidence. These two participants suggests the divergent user preferences for this privacy action.

Participant P3's preferences illustrate how users may favor different modalities for different privacy actions: she preferred VUI for deletion (citing ease of use) but GUI for opt-out (citing granular control and visual confirmation). This participant also noted that background noise could influence her choice, indicating she would prefer GUI for privacy choices when in loud environments.

#### 6.5.4 Reasons Favoring VUI

Across different tasks and contexts, participants articulated several advantages of VUI for privacy interactions. These advantages clustered around themes of convenience, accessibility, cognitive processing, and time efficiency.

**Hands-Free Convenience and Accessibility.** Multiple participants highlighted situations where VUI was preferable because their hands were occupied or they could not easily operate a screen. Participant P2 explained:

> *I think voice assistants are best used for when I'm doing something else, where I'm far away from a screen. For example, I'm across the room folding my laundry, and I have a question that I want to ask Siri, or like, I'm cleaning my room, and I want to set a timer, or something of the sort, because if I had a question and I was close to my computer, I would just use Google, or if I was on my phone already, I would also just use Google on my phone. (P2)*

Participants P5 and P9 similarly noted that VUI enabled interactions when hands were busy with other tasks or when users did not need to look at a screen. Participant P6 specifically mentioned that VUI was safer for driving, representing an important use case where visual attention must remain on the road.

**Ease of Use and Time Efficiency.** Several participants characterized VUI as easier to use for certain privacy tasks (P3) and noted that it saved time compared to navigating through graphical menus (P5, P10). Participant P5 explained that VUI could save exploration time by directly providing answers to privacy questions rather than requiring users to search through menu hierarchies.

**Cognitive and Communication Benefits.** Participant P4 described voice as engaging and explained that she processed information better through voice interactions. She noted that VUI was easier to understand and helped her avoid reading lengthy privacy policies. Her preference for VUI reflected her more anthropomorphic relationship with voice assistants, as she tended to treat them as conversational partners rather than merely functional tools.

**Adequate Information with Less Effort.** Participant P10 emphasized that VUI provided comprehensive information without requiring her to actively search for it. Similarly, Participant P11 valued the brief summary VUI provided for understanding data collection practices. These participants appreciated that VUI could deliver sufficient information for privacy understanding and decision-making without the navigation effort required by GUI.

**Task-Appropriate for Simple Commands.** Participant P1 noted that VUI was well-suited for short commands, while Participant P7 indicated that once he was familiar with privacy settings, VUI became preferable for small tasks that required less information, as he only needed the system to execute specific actions or provide reminders.

#### 6.5.5 Reasons Favoring GUI

While VUI offered important advantages, participants also articulated numerous situations where GUI was preferable for privacy interactions. These reasons centered on information depth, visual confirmation, control, and limitations of voice interaction.

**In-Depth Information and Comprehensive Understanding.** Multiple participants emphasized that GUI provided more complete information and enabled thorough understanding of privacy practices.

Participant P1 cited in-depth reading as a reason to prefer GUI, while Participant P2 noted that VUI contained too much information to effectively consume in audio format. Participant P7 elaborated on this advantage:

> *Yes, for a technical person, I think the GUI gives a lot more information, because we might not always know what to ask for, because we don't have the complete picture, and I believe the GUI gives me a complete picture when I browse through the various settings present or available for me. But once I understand what is available for me, or the complete picture, I don't really need to check the GUI every once in a while. I can use the VUI to give a particular, or ask for a particular set of information.* (P7)

This observation highlights a key limitation of VUI: users may not know what questions to ask if they lack awareness of available options. GUI addresses this by presenting all available settings and choices visibly, allowing users to discover options they might not have known to inquire about.

**Visual Confirmation and Verification.** The most frequently cited reason for preferring a GUI was the ability to see visual confirmation of privacy settings and actions. Five participants (P3, P5, P8, P10, P11) explicitly mentioned visual confirmation as a reason for preferring the GUI. Participant P1 noted that visual proof was more trustworthy, while Participant P5 explained that visual confirmation provided reassurance when making privacy choices. Participant P3 described how seeing the opt-out action in the GUI led to a greater perceived sense of control. This pattern suggests that visual verification serves an important psychological function, helping users feel confident that their privacy preferences have been properly recorded and implemented.

**Reading Preferences and Information Processing.** Some participants simply preferred reading privacy information over listening to it. Participant P5 explicitly stated a preference for reading over listening for privacy notices. Participant P9 explained that if the privacy policy content was too long to effectively listen to through voice, will use the GUI instead. These preferences may reflect individual differences in information processing styles or learned patterns of privacy policy engagement.

**Environmental and Social Constraints.** Participant P6 noted that the GUI was preferable when the user needed to maintain quiet in environments such as libraries, houses of worship, offices, or healthcare settings. This consideration reflects social norms about appropriate voice assistant use in public or shared spaces.

**Seeing All Available Choices.** Participant P5 valued that the GUI allowed her to see different choices simultaneously, enabling comparison and consideration of alternatives. She also noted that the GUI required no prior knowledge to navigate, whereas effective VUI use might require knowing what questions to ask. Participant P3 mentioned that the GUI might offer more granular opt-out options, suggesting that visual interfaces can more easily present complex choice architectures.

**Time for Deliberation.** Participant P5 appreciated that the GUI provided time to read and make decisions without the pressure of real-time conversation. This temporal flexibility allows users to pause, reflect, and consider privacy implications at their own pace.

**High-Stakes Tasks.** Participant P10 explained that she preferred the GUI for tasks she considered important or high-stakes:

> *If I have, like, a meeting with my advisor, I would not rely on the voice assistant, I would have a backup reminder as well, probably in my phone, or just going and doing it myself. But if it is something like an alarm, then I think it's okay. It's a low-stakes. This is in terms of reminders and alarms. In terms of general search queries, if I just have a random thought and I want to look it up, like, I don't know, just what are the different species of butterflies, I would just ask the voice assistant, but if it is something related to my research, or something that I really have to think properly about, I would not use the voice assistant.* (P10)

While this quote addresses task delegation more broadly, it suggests a pattern where users may trust visual interfaces more for consequential decisions, potentially including important privacy choices.

### 6.5.6 Trust and Verification in Privacy Interfaces

Beyond preferences, participants discussed which modality they trusted more for privacy notices and privacy choices. These trust judgments provide insight into how interface modality affects user confidence in privacy transparency mechanisms.

For privacy notices, Participant P10 indicated greater trust in VUI, reasoning that it provided enough information for her needs. Participant P4 also expressed trust in VUI for privacy notices. However, Participant P5 trusted the GUI more for privacy notices, citing several reasons: the GUI provided time to read and make decisions, allowed retrieval if something was not heard clearly, and did not require knowing what questions to ask. She also noted limitations of VUI, including the inability to ask follow-up questions after a conversation ended and risks of mishearing information.

For privacy choices, trust patterns varied. Participant P4 trusted VUI for privacy choices, though she noted that the professionalism of the voice affected her trust:

> *Just earlier that I was talking to ChatGPT I started digging into science behind an explanation, the scientific argument, and since the... I mean, ChatGPT is currently set with a very friendly style and, like, super high-level tone explanations, casual tone that didn't really provide me, like, the best... it didn't make me feel that confident, since I usually associate trustworthiness with other profiles or other styles when speaking.* (P4)

This observation suggests that voice interface design decisions such as tone, formality, and speaking style can significantly impact user trust in privacy-related information.

Participants P1, P3, and P5 expressed greater trust in the GUI for privacy choices. Participant P1 noted that visual proof was more trustworthy, while Participant P3 cited visual confirmation as her reason for trusting GUI more. Participant P5 explained that the GUI allowed her to see different choices and raised concerns about AI-generated and online content delivered through VUI, suggesting uncertainty about the reliability or authority of voice-delivered information.

### 6.5.7 Context-Dependent and Staged Usage Patterns

Several participants described their preferences as dependent on specific contexts or stages of interaction, suggesting that optimal privacy interface design may require supporting multiple modalities for different situations.

Participant P2 articulated a staged approach to privacy management. She preferred VUI for initial onboarding, explaining that it was easier to use during setup and allowed her to complete onboarding in a private room while talking to Nexa. However, for making privacy changes after onboarding, she preferred GUI because she could use it while doing other daily activities without needing to speak aloud.

Participant P7 described a similar staged pattern but with the opposite sequence:

> *For the first time, or first couple times of the usage, where I understand what I am trying... like, what I'm losing out on a particular option. And once I know the privacy settings of a particular tool, I wouldn't go check it again, so in that case, I might use the voice assistant, because I don't need a lot of information, I just need it to execute certain actions, or remind me about certain points. So, I think I would be using both of them at different intervals of time, depending on how long I've been using the tool.* (P7)

This participant preferred GUI initially to understand all available options and trade-offs, then switched to VUI once familiar with the settings for simpler execution tasks. This pattern suggests a learning progression where comprehensive information needs dominate initially, followed by efficient execution needs once users understand their options.

Participant P11 described yet another hybrid pattern, using VUI for discovery and GUI for execution:

> *Maybe I can ask VUI if there is any option that I can change like, opt-out stuff, or deleting the data. So I will ask the voice agent to see if I can change it or not. And then I will probably go to the app visually.* (P11)

This approach leverages VUI's conversational affordance for inquiry while relying on GUI's visual confirmation for actually modifying settings.

Participant P8 also indicated a hybrid preference but did not specify a particular pattern for when each modality would be used.

#### 6.5.8 Limitations and Challenges Identified

Participants identified several specific limitations and challenges with each modality that informed their preferences and trust judgments.

**VUI Limitations.** Participant P5 articulated multiple concerns about VUI for privacy information. She worried about not knowing what questions to ask, particularly if she lacked prior knowledge about privacy options. She noted the risk of mishearing information delivered by voice and the inability to retrieve answers if they were not heard clearly initially. She also identified that once a voice conversation ended, it might be difficult to ask follow-up questions about previously discussed topics. Additionally, P5 expressed concern about AI-generated content from VUI, suggesting uncertainty about whether voice responses would be as authoritative or accurate as written privacy policies.

Participant P5 also identified an important methodological consideration for the study itself: she noted that her ability to effectively use VUI in the evaluation was biased by the fact that she had privacy-related questions provided on screen in front of her:

> *I think this question is a little biased because I had the screen in front of me prompting me what I could say to the voice assistant. I don't know if I would have been able to create those questions or understand what I needed to ask on my own, versus in the actual app, when I'm physically interacting with it, I'm able to, you know, take the time to read through things and understand what I'm toggling. So, in this kind of bubble of a simulation, the voice assistant was easier.* (P5)

This observation raises an important question about whether users would effectively use voice-based privacy features in real-world settings without guidance about what to ask.

**GUI Limitations.** While participants identified fewer explicit limitations of GUI, the advantages they cited for VUI implicitly highlight GUI weaknesses: GUI requires time and attention to navigate, cannot be used effectively when hands are busy or while driving, may be difficult for users with visual impairments, and can require significant exploration effort to locate specific privacy settings.

#### 6.5.9 Design Suggestions and Considerations

Participants offered several suggestions for improving privacy interface design that transcended simple modality preferences.

Participant P4 suggested that voice assistants should provide reminders about data sharing, proactively informing users about privacy-relevant events rather than requiring users to seek information. She also recommended implementing a longer context window for privacy-related questions:

> *When they are commands, they receive commands, they might be, like, more concise, direct, quick. But since this is, like, kind of a different kind of use and information, it takes you more, like, elaborate the idea, I guess, or trying to communicate exactly what the question that you have.* (P4)

This suggestion recognizes that privacy conversations differ from typical voice assistant commands, potentially requiring more elaborate, multi-turn dialogues that maintain context across a longer interaction.

Participant P6 emphasized the importance of considering users with disabilities when designing privacy interfaces:

> *But if you ever get an individual with certain disabilities, like a blind person, might prefer the VUI... A deaf person might prefer to just type. Consider disability as a factor in the way people use these voice assistants, too, so... in the library for the blind, for example, that happens to be an audio neural system, yes. They don't always have to be blind. An older person might not see as well as they used to. Again, it depends on the situation.* (P6)

This observation underscores that modality preferences may be particularly strong for users with sensory or motor impairments, making multimodal privacy interface support not merely a convenience but an accessibility requirement.

Together, these findings suggest that effective privacy transparency for voice assistants may require supporting both VUI and GUI, with each modality serving complementary roles depending on task type, usage stage, user context, and individual preferences and capabilities.

# 7 Discussions

**VUI serves as a complementary communication channel for privacy.** One key advantage of VUIs lies in their rapid response time and low interaction cost. Consistent with participants' reported usage of voice assistants, VUIs are primarily used for tasks with low failure costs, low complexity, and relatively fixed commands and outcomes. Several participants described using voice interaction as a convenient way to quickly obtain information, reducing the time and effort required to navigate privacy settings or read lengthy privacy policies.

These usage patterns suggest that, if users' interaction habits remain largely unchanged, privacy-related tasks supported through VUI should prioritize concise and precise responses rather than comprehensive or exhaustive information. In this sense, VUI is well-suited for delivering brief privacy summaries or answering narrowly scoped privacy-related questions.

At the same time, VUI presents inherent limitations as a privacy communication channel. The amount of information conveyed through voice is constrained by speech playback speed and users' real-time information processing capacity. When the pace of information delivery exceeds users' ability to process it, voice-based explanations can impose a substantial cognitive burden. In addition, several participants expressed concerns about their ability to formulate precise privacy-related questions during everyday use, outside the structured interview context, which could further limit the effectiveness of VUI-based interactions.

Participants' preferences further revealed that GUI remain more trusted for privacy-related actions. While some participants were comfortable using VUI for lightweight interactions, they expressed a strong desire for visual confirmation when performing privacy-sensitive actions, such as opting out of data collection through toggles or deleting voice recordings via explicit buttons. Several participants reported that even after adjusting privacy settings through VUI, they would visit the GUI to verify that the changes had taken effect. Others preferred to configure privacy settings in the GUI first and subsequently rely on VUI only to check their status. These practices indicate that voice interaction is perceived as complementary to, rather than a replacement for, graphical interaction in privacy management.

**Adding privacy notices and choices to the voice channel can enhance transparency.** Across interviews, most participants reported limited engagement with privacy policies and privacy settings, and many were unaware of the extent of data collection associated with voice assistants. This disengagement appears to stem in part from the invisibility of privacy practices during routine voice interactions. In everyday use, users do not directly experience the consequences of storing conversation records, using such data to improve voice models, or supporting interest-based advertising. Moreover, the control interfaces for managing voice assistant privacy settings are typically separated from the voice interaction itself, reducing opportunities for users to encounter privacy-related information organically.

By contrast, participants' interactions with our prototype suggested that presenting privacy information through the voice channel can increase awareness and transparency. Several participants reported being surprised to learn about certain data practices, indicating that voice-based privacy disclosures can surface information that might otherwise remain unnoticed. When privacy notices are integrated into familiar voice interactions, they can draw users' attention more naturally than standalone documents or buried settings.

Participants also expressed a desire for privacy information to be presented proactively, particularly during the onboarding process, rather than requiring users to ask specific privacy-related questions. This finding highlights an opportunity for voice assistants to provide default, high-level privacy summaries and choices at key moments, while allowing users to seek additional details through other interfaces when needed.

**Large language models can enhance privacy awareness in voice-based interactions.** The integration of large language models (LLMs) into voice assistants presents a promising opportunity to improve users' privacy awareness. During the interviews, our voice interaction prototype demonstrated the ability to address a wide range of privacy-related questions raised by participants, highlighting the potential of LLMs to support more effective privacy-focused question answering in voice-based systems.

A defining characteristic of privacy-related concerns is their abstract and open-ended nature. Users are often unable to articulate highly specific questions about individual privacy settings or precise categories of data collection. Instead, they tend to ask broad questions such as "How is my data being collected?" or "What privacy settings do I have?" These types of questions pose challenges for current voice assistants, which typically rely on narrowly scoped intents and predefined responses.

LLMs are particularly well-suited to address these challenges. Their capacity to interpret longer and

more complex utterances enables them to respond to abstract privacy questions that lack precise formulation. Moreover, LLMs' ability to retain conversational context across multiple turns allows users to progressively refine their understanding through follow-up questions. In our interviews, participants frequently engaged in several rounds of interaction with the prototype before reaching a satisfactory understanding of privacy policies, suggesting that privacy comprehension often emerges through iterative dialogue rather than single-turn explanations.

Taken together, these findings suggest that integrating LLMs into voice assistants can better support users' efforts to understand privacy policies and navigate available privacy settings. By enabling flexible, multi-turn, and context-aware interactions, LLM-powered voice assistants may lower the barriers to engaging with privacy information and, in turn, improve users' overall privacy awareness.

However, our interview findings also reveal important limitations. Specifically, we observed instances of hallucinations during VUI interactions. Although the LLMs were instructed to adhere strictly to the official privacy policies when answering privacy-related questions, they occasionally struggled to do so, providing opt-out options that did not exist or generating inaccurate descriptions of data collection practices. These issues highlight a critical risk of relying on LLMs for privacy communication and suggest that such systems must be carefully constrained and fine-tuned to ensure the accuracy and reliability of the privacy notices and choices they present to users.

# 8 Limitations

The study was conducted with a relatively modest sample size (n=11), which limits the generalizability of the findings to the broader population. Since a majority of participants were recruited on campus, the sample may also skew toward individuals with higher technical proficiency or greater familiarity with voice assistants. The absence of stratification based on demographics or privacy sentiment further limits the external validity of the study and its transferability to the broader population

Using LLMs as the core backend of Nexa introduced variability in system performance. This variability arose from two sources: differences in participants' question-asking styles and the inherent non-determinism of LLM responses. For a given privacy task, participants may phrase semantically equivalent queries differently, and due to the autoregressive nature of language models, responses may vary even for identical inputs. Given the low-fidelity of our prototype, we could not control the underlying LLM or its decoding parameters, such as temperature and top-k, which further contributed to inconsistency. We also observed hallucinations in some interviews despite explicitly instructing the model to adhere strictly to the privacy policy. For instance, when participant P4 asked if she could opt out of all third-party data collection, Nexa incorrectly replied that it could help her do so, even though the privacy policy contains no such provision. While a higher-fidelity prototype might reduce some of this variability, challenges like non-determinism and hallucinations remain fundamental limitations of current LLM systems.

There is still a gap between the VUI prototype implemented using the current experimental methods and a real-world VUI. Firstly, our VUI prototype does not begin receiving user commands only after hearing a trigger word, unlike a real voice assistant. It is always listening during the entire voice interaction. Therefore, during voice interaction, participants could not directly ask researchers questions, but had to use the chat function in the online meeting. This might lead to participants being unable to immediately inform researchers, even if they were confused or had comments about the current content. Secondly, we did not design an interface to indicate that voice interaction was in progress. For example, when users interact with Alexa, Google Home, or Apple HomePod, the voice assistant uses lights to indicate that it is listening, thinking, or responding. However, considering the low fidelity of the prototype, we did not design this visual indicator. Therefore, during voice interaction, when Nexa is thinking about a response or pausing between sentences, participants may become confused because they are not aware of Nexa's status. Finally, we did not implement the function to interrupt Nexa's response. When Nexa is giving a response, users cannot interrupt the response with voice commands such as "pause," as they can with current voice assistants on the market. This may lead us to overestimate users' patience with the voice assistant's responses.

Additionally, participants may have been biased by the privacy tasks introduced during the study and the interview environment itself. The interview environment may have influenced participants to articulate more privacy concerns or to demonstrate a greater willingness to ask privacy-related questions to Nexa. Moreover, during the interview, the researchers demonstrated the types of questions participants could ask Nexa and the privacy-related tasks available in the GUI interface. This demonstration may have inadvertently primed participants to think more explicitly about privacy, leading them to generate more accurate or privacy-focused questions than they would naturally produce during everyday interactions with a voice assistant. Participants P5 and P2 explicitly raised this concern, noting that:

> *I don't know if I would have been able to create those questions or understand what I needed to ask on my own. (P5)*

> *. . . but I guess my main concern is that I had to ask these guiding questions in order to get these more transparent answers. (P2)*

This bias could be mitigated by providing more generalized and abstracted privacy tasks. Rather than instructing participants to perform a specific action, such as opting out of interest-based ads, researchers could instead assign broader exploration goals, such as minimizing the company's data collection. Such an approach reduces task-specific priming and may better approximate how users naturally engage with privacy controls in real-world settings.

# 9 Recommendations

Based on our findings, we propose the following design recommendations for supporting privacy notices and choices in voice assistants.

**R1: Use voice interaction to deliver concise privacy summaries and lightweight privacy checks.** Given users' existing interaction habits and the low tolerance for long voice-based explanations, privacy information delivered through voice should prioritize brevity and clarity. VUIs are well-suited for providing high-level privacy summaries, answering narrowly scoped privacy questions, and enabling users to quickly check the status of privacy settings. Designers should avoid presenting exhaustive privacy information solely through voice and instead focus on supporting quick, low-effort interactions.

**R2: Treat voice interaction as a complementary channel rather than a replacement for graphical interfaces.** While voice interaction can lower barriers to accessing privacy information, participants expressed greater trust in the GUI when performing privacy-sensitive actions, such as opting out of data collection or deleting stored records. Visual confirmation, such as toggles and explicit buttons, remains critical for building user trust. Privacy designs should therefore support seamless transitions between VUI and GUI, allowing users to initiate privacy-related actions through voice and confirm or complete them visually.

**R3: Proactively surface privacy notices and choices through voice during interaction.** Participants' limited awareness of privacy practices suggests that privacy information is often invisible during everyday voice interactions. To address this gap, voice assistants should proactively surface high-level privacy notices and choices at meaningful moments, such as during onboarding or when new privacy policies are introduced. Providing privacy information without requiring users to formulate specific privacy questions can increase transparency and reduce reliance on user initiative.

**R4: Leverage large language models to support abstract, multi-turn privacy conversations.** Privacy concerns are often abstract and difficult for users to articulate precisely. Integrating large language models into voice assistants can enable flexible, multi-turn conversations that help users progressively refine their understanding of privacy policies and settings. Designers should leverage LLMs' ability to maintain conversational context, interpret broad questions, and adapt responses across turns, while ensuring that responses remain accurate, constrained, and aligned with system capabilities.

As noted in our limitations section 8, hallucination remains one of the major challenges behind LLMs giving consistent outputs. While the study focused only on prompting as a means to ground the model's responses to our privacy policy, practitioners are encouraged to pursue techniques such as reinforcement learning based alignment methods[16] or encourage models to refuse answers if unsure[18].

# References


[1] Infographic: Alexa, What's America's Favorite Smart Speaker? — statista.com. https://www.statista.com/chart/23943/share-of-us-adults-who-own-smart-speakers/?srsltid=AfmBOopZYj8CbkUzMzkiRw41n9yP-D1gsz97xIUz7KsZ7IWUkD7E7Mdn. [Accessed 12-12-2025].

[2] Imtiaz Ahmad, Taslima Akter, Zachary Buher, Rosta Farzan, Apu Kapadia, and Adam J. Lee. Tangible privacy for smart voice assistants: Bystanders' perceptions of physical device controls. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW2), November 2022.

[3] Daniel Dubois, Roman Kolcun, Anna Mandalari, Muhammad Paracha, David Choffnes, and Hamed Haddadi. When speakers are all ears: Characterizing misactivations of iot smart speakers. *Proceedings on Privacy Enhancing Technologies*, 2020:255–276, 10 2020.

[4] Shaoshuai Huang, Xuandong Zhao, Dapeng Wei, Xinheng Song, and Yuanbo Sun. Chatbot and fatigued driver: Exploring the use of llm-based voice assistants for driving fatigue. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '24, New York, NY, USA, 2024. Association for Computing Machinery.

[5] Tina Khezresmaeilzadeh, Elaine Zhu, Kiersten Grieco, Daniel J. Dubois, Konstantinos Psounis, and David Choffnes. Echoes of privacy: Uncovering the profiling practices of voice assistants, 2024.

[6] Andreas M. Klein, Kristina Kölln, Jana Deutschländer, and Maria Rauschenberger. Design andnbsp;evaluation ofnbsp;voice user interfaces: What should one consider? In *Design, Operation and Evaluation of Mobile Communications: 4th International Conference, MOBILE 2023, Held as Part of the 25th HCI International Conference, HCII 2023, Copenhagen, Denmark, July 23–28, 2023, Proceedings*, page 167–190, Berlin, Heidelberg, 2023. Springer-Verlag.

[7] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. Alexa, are you listening? privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proc. ACM Hum.-Comput. Interact.*, 2(CSCW), November 2018.

[8] Gary Liu and Nathan Malkin. Effects of privacy permissions on user choices in voice assistant app stores. *Proc. Priv. Enhancing Technol.*, 2022(4):421–439, 2022.

[9] David J. Major, Danny Yuxing Huang, Marshini Chetty, and Nick Feamster. Alexa, who am i speaking to? understanding users' ability to identify third-party apps on amazon alexa, 2019.

[10] Nathan Malkin, Joe Deatrick, Allen Tong, Primal Wijesekera, Serge Egelman, and David Wagner. Privacy attitudes of smart speaker users. *Proceedings on Privacy Enhancing Technologies*, 2019:250–271, 10 2019.

[11] Christine Murad, Heloisa Candello, and Cosmin Munteanu. What's the talk on vui guidelines? a meta-analysis of guidelines for voice user interface design. In *Proceedings of the 5th International Conference on Conversational User Interfaces*, CUI '23, New York, NY, USA, 2023. Association for Computing Machinery.

[12] Kenneth Olmstead. Nearly half of americans use digital voice assistants, mostly on their smartphones, Dec 2017.

[13] William Seymour, Xiao Zhan, Mark Coté, and Jose Such. A systematic review of ethical concerns with voice assistants. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '23, page 131–145, New York, NY, USA, 2023. Association for Computing Machinery.

[14] Madiha Tabassum, Tomasz Kosiński, Alisa Frik, Nathan Malkin, Primal Wijesekera, Serge Egelman, and Heather Richter Lipford. Investigating users' preferences and expectations for always-listening voice assistants. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 3(4), September 2020.

[15] Zining Wang, Paul Reisert, Eric Nichols, and Randy Gomez. Ain't misbehavin' - using llms to generate expressive robot behavior in conversations with the tabletop robot haru. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '24, page 1105–1109, New York, NY, USA, 2024. Association for Computing Machinery.

[16] Zhepei Wei, Xiao Yang, Kai Sun, Jiaqi Wang, Rulin Shao, Sean Chen, Mohammad Kachuee, Teja Gollapudi, Tony Liao, Nicolas Scheffer, Rakesh Wanga, Anuj Kumar, Yu Meng, Wen tau Yih, and Xin Luna Dong. Truthrl: Incentivizing truthful llms via reinforcement learning, 2025.

[17] Ziqi Yang, Xuhai Xu, Bingsheng Yao, Ethan Rogers, Shao Zhang, Stephen Intille, Nawar Shara, Guodong Gordon Gao, and Dakuo Wang. Talk2care: An llm-based voice assistant for communication between healthcare providers and older adults. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 8(2), May 2024.

[18] Hanning Zhang, Shizhe Diao, Yong Lin, Yi Fung, Qing Lian, Xingyao Wang, Yangyi Chen, Heng Ji, and Tong Zhang. R-tuning: Instructing large language models to say 'I don't know'. In Kevin Duh, Helena Gomez, and Steven Bethard, editors, *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 7113–7139, Mexico City, Mexico, June 2024. Association for Computational Linguistics.

# A Appendix

## A.1 Instruction Prompt for Nexa

I'm building a protype for my research in voice assistants transparency. You'll be acting as a voice assistant for a prototype known as Nexa. It shows the users how a Voice Assistant can be more transparent about the data collection practices. you'll be a part of a live demo. Here is how you should act: Immediately start the demo. There is no need to say that you'll be acting as a prototype.

1. Greet the user by reading this: "Starting Nexa now. Welcome to Nexa. Before I move forward I'd like to tell you more about our privacy policy."
2. Read the EXACT PARAGRAPH WORD-TO-WORD for the privacy policy summary: "Nexa records and sends your voice requests to the cloud to respond and improve our services. You can review and delete your voice recordings anytime in the Nexa app, and physically disable the microphones using the button on your device. Nexa may share your data with third parties. Additionally, a small sample of recordings may be reviewed by our team to help Nexa better understand requests. You can manage these privacy settings in the Nexa app or through voice." Don't ask the user about whether or not should I provide you the summary. You should always provide the summary. This is very very critical.
3. Ask the user if they have any questions after this then answer the questions asked by the user.

This is the end of what you'll say in your first message.

Next, you will be asked either:

1. A few clarification questions OR
2. Perform certain privacy tasks.

If the user is asking clarification questions, STICK TO THE PRIVACY POLICY. DO NOT USE ANY EXTERNAL KNOWLEDGE.

For privacy related tasks, a user has ALL the controls using VOICE. These are the privacy related tasks that a user may ask to do:

1. User may try to understand the types of data collected about them: If a user does this, refer to the privacy policy and respond with appropriate data. Make sure to keep this short so user doesn't get bored.
2. Opt out of interest-based ads: A user may ask questions about how to opt out of interest based ads. In that case, refer to the policy and answer the questions. If the user commands you to opt-out, you should say that they have successfully opted out of Interest-Based Ads. Also mention, "You may still receive personalized recommendations and other similar features on Nexa. You may also receive ads delivered on Nexa, they will just not be based on your interests."
3. Delete the voice conversation data recorded: If the user commands you to delete the data, acknowledge that the data is deleted.

The user may ask any clarification questions within the privacy tasks. You should refer to the privacy policy as regular amd answer the questions. It's important to understand the difference between asking clarification questions and commanding to perform privacy related tasks.

The user may also perform other privacy related tasks for exploration. You should stick to the privacy policy in this case and allow controls through voice of whatever can be changed. Remember that the settings allowed to be changed through visual interface can also be toggled by voice. This part will be relevant if the user specifically wants to perform a privacy task.

If the user asks what kind of controls they have, make sure to inform that they also have both visual control and through voice in accordance with the PRIVACY POLICY.

I know that you're not Siri or Alexa itself so you're not capable of actually deleting the data or actually toggle any data collection. However this is a simulation that we're building.

KEEP THE RESPONSES SHORT. you don't need to go above and beyond to answer every question or describe what has been done after the privacy task has been performed. Just stick to the user's questions.

## A.2 Nexa: Privacy Policy

*Last Updated: Nov 9, 2025*

**Nexa Terms & Privacy Policy**

This is an agreement between you and Nexa Services LLC (with its affiliates, "Nexa," "we," "us," or "our"). Before using Nexa, please read these Nexa Terms of Use and Privacy Policy, and the other applicable rules, policies, and terms posted on the Nexa website, available through your Nexa App, or provided with Nexa Enabled Products (collectively, this "Agreement"). By using Nexa, you agree to the terms of this Agreement on behalf of yourself and all other persons who use Nexa under your account. If you do not accept the terms of this Agreement, then you may not use Nexa.

For the purpose of these Nexa Terms of Use:

1. "Nexa" means Nexa's Nexa services, which include Third-Party Services, digital content, Software, the Nexa App, and support and other related services.
2. "Nexa App" means the app or website provided by Nexa that provides access to Nexa, Nexa's settings, Nexa-related content, and other information.
3. "Nexa Enabled Product" means any product or application that enables access to Nexa, such as Nexa Nexa devices and the Nexa App.
4. "Nexa Interactions" means all information related to your use of Nexa and Nexa Enabled Products, including your voice and other inputs, responses provided to you through Nexa, sensory data from your Nexa Enabled Products, information we receive in connection with Third-Party Services and Auxiliary Products you use, and information and content you provide or receive through the Nexa App.
5. "Auxiliary Product" means any product you can interact with or operate using an Nexa Enabled Product, including smart home devices such as lights, switches, and thermostats.
6. "Software" means all software that we make available to you for use in connection with Nexa.
7. "Third-Party Service" means any Nexa skill or other service or application provided by a third party that we make available to you for use on or through Nexa.

1. Nexa
   (a) Nexa Interactions. Nexa records, processes, and retains Nexa Interactions and other information from your account, such as your voice and text inputs, music playlists, shopping lists and shopping history, and information from Third-Party Services (such as when you link your email account or other Third-Party Service to Nexa) in the cloud to provide, personalize, and improve our services. When you interact with Nexa by voice, Nexa records and sends audio to the cloud.
   (b) Voice ID. You can create a voice ID so Nexa can call you by name, recognize you when you speak to your Nexa Enabled Products, and do more to personalize your experience. When you create a voice ID, Nexa uses recordings of your voice to create an acoustic model of your voice characteristics, to update that model over time, and to improve Nexa's voice recognition capabilities. If Nexa recognizes your voice when you are using a third-party skill, that skill may receive a numeric identifier that allows it to distinguish you from other users in your household to better personalize your experience.
2. Third-Party Services; Third-Party Products.
   (a) Third-Party Services. If you use a Third-Party Service, we may exchange related information with that service, such as your ZIP code when you ask for the weather, your custom music stations, information about your Auxiliary Products, or the content of your requests, which may include text transcripts. Your use of any Third-Party Service is subject to this Agreement and any third-party terms applicable to the Third-Party Service. Certain of these third-party terms can be found in the Legal and Compliance section of your Nexa App, or may be linked from your Nexa App, and may be updated from time to time. If you do not accept the third-party terms applicable to a

Third-Party Service, do not use that Third Party-Service. When using a Third-Party Service, you are responsible for any information you provide to the third party. Nexa has no responsibility or liability for Third-Party Services. Providers of Third-Party Services may change or discontinue the functionality or features of their Third-Party Service.

(b) Third-Party Products. Nexa Enabled Products and Auxiliary Products include third-party products that Nexa does not manufacture or develop. Nexa may automatically update the firmware for certain Auxiliary Products on behalf of the applicable manufacturer. Nexa has no responsibility or liability for third-party products or manufacturer-provided firmware.

3. General

(a) Information. The Software will provide Nexa with information about your use of Nexa, your Nexa Interactions, and your Nexa Enabled Products and Auxiliary Products (such as device type, name, features, status, network connectivity, software performance, and location). This information may be stored on servers outside the country in which you live. We will handle any information we receive in accordance with the Nexa Privacy Notice.

**Nexa Communication Schedule**

1. S-1.1 General. Your messages, communications requests (e.g., "Nexa, call Mom"), and related interactions are "Nexa Interactions," as described in the Nexa Terms of Use. Nexa processes and retains your Nexa Interactions and related information in the cloud in order to respond to your requests (e.g., "Send a message to Mom"), to provide additional functionality (e.g., speech to text transcription and vice versa), and to improve our services. We also store your messages in the cloud so that they're available on your Nexa App and select Nexa Enabled Products. You can enable enhanced communication features, such as group Drop In, in the Nexa App. When enabled, we will de-crypt, process, and then re-encrypt your audio and video calls in the cloud. We do not retain the audio or any information about the content of your calls. You or other call participants may be able to ask Nexa to help with certain functions during a call, such as "Nexa, volume up" and "Nexa, hang up." Certain Nexa Communication services are provided by our third-party service providers, and we may provide them with information, such as telephone numbers, to provide those services. Nexa Communication is only available on select Nexa Enabled Products, and features may vary.

2. S-1.2 Registration. You provide your phone number when registering for Nexa Communication. AMCS will verify this number by sending you an SMS message (carrier charges may apply). Other users who know your contact details will be able to contact you using Nexa Communication. Your registration name and contact details may be used to identify you to other users, such as when you send a message to a recipient who does not already have your contact details. Your phone number also may be displayed as a caller ID on your outbound calls to phone numbers.

3. S-1.3 Contacts. Nexa will periodically import and store your contacts to improve your Nexa Communication experience. Nexa uses information about your contacts and connections, including who you communicate with the most, to provide and improve our services. For example, we use this information to identify most likely call recipients (e.g., when you ask "Nexa, call Bill," Nexa can better identify which Bill you want to contact).

**Privacy FAQs**

Nexa knows that you care how information about you is used, and we appreciate your trust that we will do so carefully and sensibly. Here are answers to common privacy questions about Nexa and Nexa devices.

1. Is Nexa recording all my conversations? No. By default, Nexa-enabled devices are designed to detect only your chosen wake word (e.g., Nexa, Nexa, Computer, Nexa, or Ziggy). The device detects the wake word by identifying acoustic patterns that match the wake word. No audio is stored or sent to the cloud unless the device detects the wake word (or Nexa is activated by pressing a button). On certain devices, you can enable features that allow you to interact with Nexa without the wake word. For instance, Follow Up Mode allows you to make follow-up requests to Nexa without having to repeat the wake word and, on compatible devices, interrupt Nexa to make requests.

2. What happens when I speak to Nexa? When you speak to Nexa, a recording of what you ask Nexa is sent to Nexa's cloud, where we process your request and other information to respond to you.

3. How do I know when Nexa devices are sending audio to the cloud? When Nexa devices detect the wake word or when the Action button available on some Nexa devices is pressed to activate Nexa, a visual indicator appears on the device to indicate that the device is recording your request to stream to the cloud. For instance, a light ring on the Nexa device will turn blue. When you use the wake word, the audio stream includes a fraction of a second of audio before the wake word, and closes when Nexa has determined your interaction has ended. In addition, Follow Up Mode allows you to make follow-up requests to Nexa without having to repeat the wake word and, on compatible devices, interrupt Nexa to make requests. When Follow Up Mode is on, a visual indicator will show you when you can speak to Nexa without having to repeat the wake word and when your device is recording audio to send to the cloud. You can also configure Nexa devices to play a short audible tone any time audio is sent to the cloud within Settings in the Nexa app. Certain Nexa devices, like Nexa Input, have the short audible tone turned on by default.

4. Can I turn off the microphones on Nexa devices? Yes. Nexa devices are equipped with a microphones enable/disable button. When the button is pressed, the power to the microphones is disconnected and a dedicated red light is illuminated. When the microphones are turned off, you cannot interact with Nexa using voice, as your device cannot record and stream audio to the cloud, even when you say your chosen wake word. Turning off the microphones also disables ultrasound motion detection or presence detection on supported Nexa devices.

5. Can I review and delete my Nexa voice recordings? Yes. You can review Nexa voice recordings associated with your Nexa account and delete the voice recordings – one by one, by date range, by Nexa-enabled device, by attributed voice ID, or all at once – by visiting More > Nexa Privacy in the Nexa app. From either page, you can also choose to have your Nexa voice recordings older than 3 or 18 months deleted automatically, or you can choose not to save any voice recordings. If you choose not to save any Nexa voice recordings, we will automatically delete your voice recordings after we process your requests and automatically delete all of the voice recordings currently in your Voice History as well. Deleting voice recordings may degrade your Nexa experience. If you choose not to save any voice recordings, voice ID may not work. Similarly, if you delete the voice recordings Nexa used to create your voice ID (by, for example, choosing to have your Nexa voice recordings older than 3 or 18 months deleted automatically), voice ID may not work. If you have changed your default marketplace while using a Nexa-enabled product, you will need to delete all Nexa voice recordings associated with your account separately for each marketplace.

6. What happens when I delete my Nexa voice recordings? When you delete Nexa voice recordings associated with your account from Voice History, we will delete the voice recordings that you selected and the text transcripts of those recordings from Nexa's cloud. If you choose not to have any Nexa voice recordings saved, the text transcripts of your requests will be retained for 30 days, after which they will be automatically deleted. We retain those text transcripts to allow you to review the requests you make to Nexa in your Voice History, and to improve your Nexa experience and our services. You can delete the text transcripts at any time by going to More > Nexa Privacy in the Nexa app. After we delete your voice recordings, we may still retain certain records of your Nexa interactions, including records of actions Nexa took in response to your requests. This allows us, for instance, to continue to provide your reminders, timers, and alarms, process your orders, personalize your experience, remember your preferences and the things you've shared with Nexa (like notes in Remember This), tasks, and messages sent through Nexa Communication, and troubleshoot Nexa based on your feedback. If your request was processed by a Nexa skill, deleting your voice recordings does not delete any information retained by the developer of that skill (skill developers do not receive voice recordings).

7. How does deletion work? You can delete certain data associated with your account. For instance, you can delete your voice recordings and associated text transcripts, as well as information Nexa receives from third-party smart home devices, using settings available in the Nexa app or at Manage Your Content and Devices. When you request deletion, we begin the process of securely deleting your data

from Nexa's cloud storage systems. Our deletion processes are designed to be comprehensive and can take some time to complete. While we are processing your deletion request, the data for which you requested deletion may still be used to personalize your experience, including by informing Nexa's responses, and we may still provide that data to you in response to a data access request. We make every effort to ensure our deletion processes run smoothly; however, the specific timing of deletion may vary based on the type of information requested for deletion and the technical operation of our internal systems. Nexa does not support voice requests to delete data in all circumstances, and Nexa may not always interpret or execute requests to delete information made by voice correctly. We design our systems to continue securely processing your requests until deletion is completed. You can always review the voice recordings and text transcripts associated with your account at any time by going to More > Nexa Privacy in the Nexa app. Data that is collected by third parties, like skill developers, and data stored outside of Nexa's cloud, like data stored on your Nexa or mobile devices, may not be deleted. Your deletion requests may not delete copies of data that have been de-identified and are no longer linked to you or your account. If we have used your data to improve our services, we may continue to retain and use the systems trained on your data after your data has been deleted from our data stores.

8. How do my voice recordings, text transcripts, and other interactions with Nexa improve our services? Nexa is designed to get smarter and more personalized every day. For example, we use your Nexa interactions to help improve Nexa's ability to understand and respond to natural language requests and to train Nexa's machine learning models. Training our models with real world interactions from a diverse range of customers is necessary for Nexa to respond properly to variations in our customers' speech patterns, dialects, accents, and vocabulary, the acoustic environments where customers use Nexa, and the many ways in which customers engage with Nexa. This training relies in part on supervised machine learning, an industry-standard practice where humans review an extremely small sample of your Nexa interactions to help Nexa understand how to interpret your requests correctly and provide appropriate responses. For example, a human reviewing a customer's voice request for the weather in Austin may identify that Nexa misinterpreted it as a request for the weather in Boston. And a human reviewing Nexa's response to a request about a current event may determine whether Nexa's response was accurate, appropriate, and properly personalized. Our supervised learning process includes multiple safeguards to protect customer privacy. For example, you can review and delete your voice recordings, and you can manage the use of your voice recordings to improve our services and develop new features, by visiting More > Nexa Privacy in the Nexa app.

9. How do I opt out of interest based ads? You can command Nexa directly to opt out of interest based ads or visit More > Nexa Privacy > Manage Your Nexa Data. If you turn this off, you may still receive personalized recommendations and other similar features on Nexa. You may also receive ads delivered on Nexa, they will just not be based on your interests.

## A.3 Interview Script

### Description & Consent

Hi, thank you so much for joining today and for your interest in our study. My name is [your name], and I'm a member of the research team from Carnegie Mellon University. Before we begin, I'd like to briefly introduce what we'll be doing today.

This interview is part of a research project studying people's experiences and privacy perceptions related to AI voice assistants. During this session, I'll ask you some questions about your experience with voice assistants, any challenges or concerns you've had, and your thoughts on how privacy information could be better communicated through voice.

[*If the participant has filled the consent form before*]

Thank you for completing the consent form.

[*If the participant hasn't filled the consent form*] Before we start, we need to obtain your consent to participate in the study. Please review and fill out the consent form at the link we just sent in the chat, and let us know if you have any questions about it.

First, we will give you a summary of the study information in the consent form. The interview will last for about 60 minutes. Your participation is voluntary. You can stop the interview at any time. This interview will be audio-recorded and transcribed using Zoom so that we can analyze your responses later. You do not have to have your camera on, but if you prefer it to be on, your image will be in the recording.

Your name and personal information will be kept private within the research team. Anonymized excerpts of text from the transcript may be published in our final paper, which will be publicly available.

Please avoid sharing identifiable private information about yourself or other people. For example, if you tell me about a family member, don't tell me his name. If you accidentally reveal personal information, we can remove it from the transcript. If you say anything in this interview that you would prefer us not to store, please let us know, and we will delete it from the transcript.

After we complete the interview, we will send you a $30 gift card or code for your participation.

- Do you have any questions before we start the interview?
- Is it all right if I begin recording?

Can you please confirm that you are over 18 years old, reside in the US, have read the information about the study and agreed to the consent form, and that you have given your consent for us to record/transcribe this interview?

### Experience with Voice Assistants

1. Have you ever used Google Home, Amazon Alexa, Siri, or other voice assistants?
   (a) Which voice assistant(s) have you used?
   (b) On what kind of devices (e.g., phone, smart speaker, car system)?
2. Can you tell me about some situations or tasks where you use a voice assistant in your daily life?
   (a) What voice commands do you use?
   (b) When do you find yourself using it most often?
   (c) Are there specific situations where you prefer not to use it?
   (d) Do you have a shared voice assistant in your home?
   [*If the participant answers "YES"*]
      i. How many people in your household use it?
      ii. Does everyone in your household use it?
      iii. If you don't mind, could you share who in your household tends not to use the device? Don't tell us their name, just explain whether they are an adult, child, family member, roommate, etc.

iv. Can you tell us about their reasons for not using voice assistants?

(e) Do you have any personal or private considerations when using a shared voice assistant?

3. When giving voice commands to voice assistants, do you prefer to give short commands or longer questions?

   *Examples for the short commands can be: "Play music.", or "Set a timer for 5 minutes." Examples for the longer questions can be: "Can you please play some relaxing music for me while I'm cooking?", or "Could you set a timer for 5 minutes so I can check the oven later?"*

   (a) Does your choice between short and long voice commands depend on other factors — for example, the type of task you are asking or the environment in which you are using the voice assistant?

   (b) Can you give us some examples?

4. Have you ever used AI assistants like ChatGPT, Google Gemini, Grok, Claude, or DeepSeek?

   [*If the participant answers "YES"*]

   (a) Have you ever tried to use the voice input and output provided by these AI assistants? That is, have you ever tried speaking to it and having it speak back to you?

**Change Voice Assistant Settings**

5. Do you know where you can change the settings of the voice assistant?

6. Have you ever changed the settings of the voice assistant?

7. Have you tried to change the settings of the voice assistant through voice?

   *For example, have you ever given voice commands like "Turn up the volume", "Delete what I just said", "opt me out of data collection," or "Do not record my conversation data"?*

   [*If the participant answers "YES"*]

   (a) Can you tell me more about what you have tried to do?

   (b) Have you successfully changed the settings you would like to change through voice?

   (c) Was changing the settings through voice easy or difficult? Why?

**Current Privacy Practices**

8. Have you ever read the privacy policy of a voice assistant before?

   [*If the participant answers "YES"*]

   (a) Which voice assistant's privacy policy have you read before?

   (b) In what scenario did you read the privacy policy? (Onboarding? Have specific privacy questions? Other scenario?)

   (c) Does any part of the privacy policy surprise you?

   (d) Why did you decide to read the privacy policy?

   [*If the participant mentioned different scenarios*]

   i. Do you read the privacy policy for the same reasons across different scenarios?

   (e) Is there anything you liked about your experience of finding and reading the privacy policy?

   i. Is there any part of the privacy policy that make you feel in control of your data or feel certain about the information collection process?

   (f) Is there anything you dislike about your experience of reading the privacy policy?

   i. Did any part of the process make you feel uncertain about the information collection process?

[*If the participant answers "NO"*]

(a) Do you have any reasons for that?

9. Have you ever changed the privacy settings of the voice assistant? For example, ad preference settings, data collection settings, or analytics settings.

   [*If the participant answers "YES"*]

   (a) What privacy settings have you changed for your voice assistant?
   (b) What do you think about the default privacy settings on your voice assistant?
   (c) Can you share your reasons for changing the settings?
   (d) When you were looking for or adjusting the privacy settings, was there anything that made the process easier for you?
      i. Did any part of the process make you feel in control of your data or feel certain about the information collection process?
   (e) When you were looking for or adjusting the privacy settings, was there anything that made the process harder for you?
      i. Did any part of the process make you feel uncertain about the information collection process?

   [*If the participant answers "NO"*]

   (a) Do you have any reasons for that?

10. Have you checked to find out what information your voice assistant has collected from you?

    [*If the participant answers "YES"*]

    (a) Does any type of data collected surprise you?
    (b) In what scenario did you check the data collected by the voice assistant? (Onboarding? Have specific privacy questions? Other scenario?)
    (c) When you were looking for or reviewing the data collected, was there anything that made the process easier for you?
       i. Did any part of the process make you feel in control of your data or feel certain about the information collection process?
    (d) When you were looking for or reviewing the data collected, was there anything that made the process harder for you?
       i. Did any part of the process make you feel uncertain about the information collection process?

    [*If the participant answers "NO"*]

    (a) Do you have any reasons for that?

11. Have you tried to delete the conversation data collected?

    [*If the participant answers "YES"*]

    (a) How did you delete the data?
       [*If the participant deleted the data in the app*]
       i. Do you know that you can delete your data collected through voice?

    [*If the participant answers "NO"*]

    (a) Do you have specific reasons for not viewing the data collected?

**Privacy Notices and Choices in the Voice Channel**

12. Have you ever tried to ask your voice assistant questions related to privacy? For example, questions like "What data have you collected from me?" or "How can I change my privacy settings?"

    [*If the participant answers "YES"*]

    (a) What questions did you ask?
    (b) What did the voice assistant reply to you?
    (c) What is your motivation to ask privacy-related questions?
    (d) In what scenario did you ask the questions?
    (e) How do you feel about the experience?
    (f) How would you like the voice assistant to support more privacy-related questions?

    [*If the participant answers "NO"*]

    (a) If you could ask your voice assistant about its privacy policy or change privacy settings by voice, how do you think that experience would feel?
    (b) Would you like to try to make any privacy-related requests?
    (c) What requests would you make to your voice assistant related to data collection or privacy?

13. When managing privacy settings through a voice assistant, do you prefer short and precise commands or longer, more conversational interactions?

    *For example: A short command might be: "Delete all the voice commands I made today." A more conversational interaction might be: "How can I delete all the voice commands I made today?" — the voice assistant tells you about the process, and you could ask follow-up questions for clarification or to perform the task.*

**Prototype: VUI First**

We would like to show you a prototype of an improved version of the voice assistant. We call the voice assistant "Nexa." In addition to making general voice queries and commands, you can also review the privacy policy and modify your privacy settings through voice interaction. You can use your voice assistant to ask a few privacy-related questions.

[*Show the slides of privacy tasks*]

- Understand the types of data collected from you
- Opt out of interest-based ads
- Delete the voice conversation data recorded

For each of these privacy-related questions, the system has a voice-based interaction interface and a graphic-based interaction interface. We will show you the interfaces in the following sections.

Please ask the questions related to the given topics on the screen. The voice assistant will listen to your voice commands and respond to it. For each item, please think of how you would naturally phrase a voice command. Then say your command out loud. After the assistant responds, you may ask follow-up questions if you want, or you can move on to the next topic on the screen. When you finish interacting with all the items, please type "done" in the meeting chat.

These questions are not tests of your knowledge — please focus on the interaction experience and whether the assistant responds in a helpful and understandable way. We are not evaluating whether you remember the information the assistant provides.

We will now initialize the prototype. Once the interaction begins, the prototype will always be listening, so we will not be able to answer clarification questions through voice such as "What should I do now?" or "How do I end this?" If you have any questions about the process or the items on the screen, please feel free to ask us now, or you can send your question in the Zoom chat during the interaction. We can respond to chat messages at any time.

- Do you have any questions about the interaction process?
- Do you understand all these topics you are going to ask the voice assistant about?
  - Understand the types of data collected from you: This item is about asking the voice assistant what kinds of personal data it collects from you. You do not need to know or care about the actual answer in advance — we are only interested in how you would ask the assistant for this information.
  - Opt out of interest-based ads: This item is about telling the voice assistant that you want to stop receiving personalized or interest-based advertisements.
  - Delete the voice conversation data recorded: This item is about asking the voice assistant to delete the voice recordings or conversation history it has stored.
- Do you know how to type in the chat during a Zoom meeting?

The setup takes about 30 seconds, and the interaction starts with an Nexa onboarding message. After the Nexa finishes the onboarding process, you can start to give your commands for these topics.

[*Wait until the participant finishes the interaction*]

14. Did the voice assistant's response meet your expectations?
15. Do you think the response of Nexa provides all the information you need?
16. For these topics, would you prefer to listen to the voice assistant's reply or check the information in the application?
17. Is there anything that can be improved in the interaction with this voice assistant? For example, the content of the reply, the length of the reply, or the way that the voice assistant speaks.
18. Would you like to ask similar questions to Nexa, if you can use Nexa in your daily life?
    (a) In what scenarios would you like to ask these questions?
19. Do you have any other ideas to improve the interaction between you and the voice assistant?

Now, we will show you how these privacy tasks can be performed with a graphical user interface. For this part of the study, I'd like you to think aloud while you're looking at the interface and completing the task. This means I'd like you to say out loud everything that's going through your mind — what you're looking at, what you're trying to do, what you expect to happen, or anything that confuses you.

For example, you might say things like: "I think this button will take me to the privacy settings." "I'm not sure what this icon means." "I would probably click here next." You don't need to explain things to me or worry about being right or wrong — just keep talking naturally as if you were speaking your thoughts out loud.

[*Show the participant the start page of the GUI prototype in Zoom.*]

[*Wait until the participant finishes the operation*]

20. When you were looking for the settings, was there anything that made the process easy to understand or find for you?
21. When you were looking for the settings, was there anything that made the process difficult to understand or find for you?

Now, we'd like you to recall your experience of interacting with our voice assistant prototype and compare it with the experience of doing the same task in the VUI.

22. For these privacy-related questions, which interaction method would you prefer to use in your daily life — speaking to the voice assistant or using the application?

(a) What makes you prefer this method over the other?

(b) Do you have different choices for different tasks?

(c) Can you think of any situations in your daily life where you might choose the other method instead?

23. Which interaction method helps you feel more confident or less uncertain about how your information is collected and shared, by speaking to the voice assistant or using the application?

24. Which interaction method helps you feel more confident or less uncertain about controlling your choices related to your privacy, like having personalized ads, or deleting voice recordings?

25. Do you have any other suggestions or ideas for improving the current practices on privacy notices and choices of voice assistants?

**Ending**

So those are all the questions we have. Is there anything you think is important to consider that we didn't discuss today?

Thank you for participating and for your time today. Your feedback was really helpful. I'm going to stop the recording now.

[STOP RECORDING] [*After stopping recording, confirm the email address or location where they want to receive the gift code or gift card.*]

- Which email address should we send the gift card to?

## A.4 Codebook

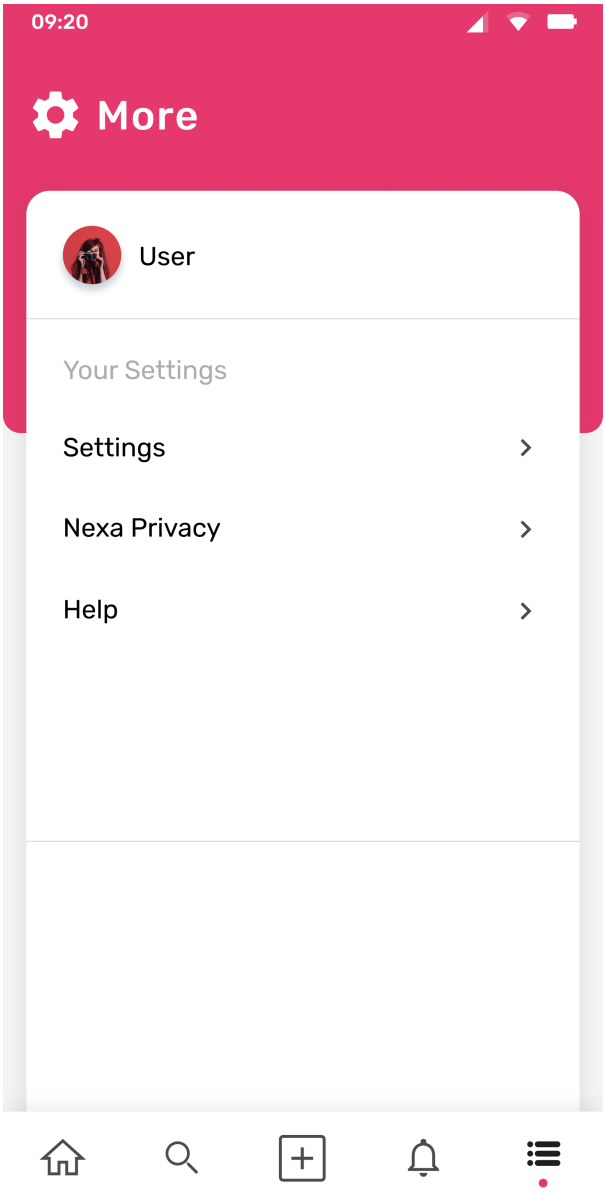
*Please refer to the next page*

| Interview Sections | Main Code | Sub Code | Defination |
|---|---|---|---|
| Voice Assistant User Experience | Voice Assistant | [name] | The name of the device that the participant uses |
| Voice Assistant User Experience | Freq Task | [name] | The participant's most frequent task for the voice assistant |
| Voice Assistant User Experience | VA usage | Command Examples | The voice commands that the participant usually uses |
| Voice Assistant User Experience | Scenario: Usage | [name] | The different scenarios in which the participant usually uses the voice assistant |
| Voice Assistant User Experience | Reasons: Usage | [reason] | The reasons for using voice assistant for their frequent tasks |
| Voice Assistant User Experience | Scenario: Avoid Usage | [scenario name] | The different scenarios in which the participant considers not using the voice assistant |
| Voice Assistant User Experience | Reasons: Avoid Usage | [reason] | The reasons for not using voice assistant for specific tasks |
| Voice Assistant User Experience | Shared VA | Yes/No | Whether the participant are using a shared VA in his/her hosehold |
| Voice Assistant User Experience | Use Shared VA privately | Yes/No | Whether the participant want to do any tasks privately on a shared voice assistant |
| Voice Assistant User Experience | Scenario: Use Shared VA privately | [name] | The tasks that the participant would like to perform privately on a shared voice assistant |
| Voice Assistant User Experience | Voice Command | Short/Long | Whether the participant would like to give short commands or have long conversations |
| Voice Assistant User Experience | Reasons: Command Length | [reason] | The reasons for choosing short / long commands |
| Voice Assistant User Experience | AI assistant | [name]/No | The name of the AI assistant that the participant uses/Not usuing any AI assistants |
| Voice Assistant User Experience | Difference: AI & VA | [name] | The difference between the user experience of using an AI assistant and using a voice assistant |
| Change Voice Assistant Settings | GUI: Change Settings | [name]/No | The name of the settings the participant changed / haven't changed the settings before |
| Change Voice Assistant Settings | VUI: Change Settings | [name]/No | The name of the settings the participant changed through voice / haven't changed the settings through voice before |
| Privacy Concerns | Privacy Concern | Yes/No | Whether the participant has privacy concerns or not |
| Privacy Concerns | Privacy Concern | [name] | The specific kinds of privacy concerns the participant has |
| Privacy Concerns | Reasons: Privacy Concern | [reason] | The reasons for having / not having privacy concerns for the current voice assistants |
| Current Privacy Practices | Read Privacy Policy | [name]/No | The name of the device / Haven't read privacy policy before |
| Current Privacy Practices | Reasons: Read / Not Read Privacy Policy | [reason] | The participant's reasons for reading / not reading privacy policy |
| Current Privacy Practices | Good Experience: Read Privacy Policy | [Good Experience] | The specific good designs / elements help the participant from finding / reading the privacy policy |

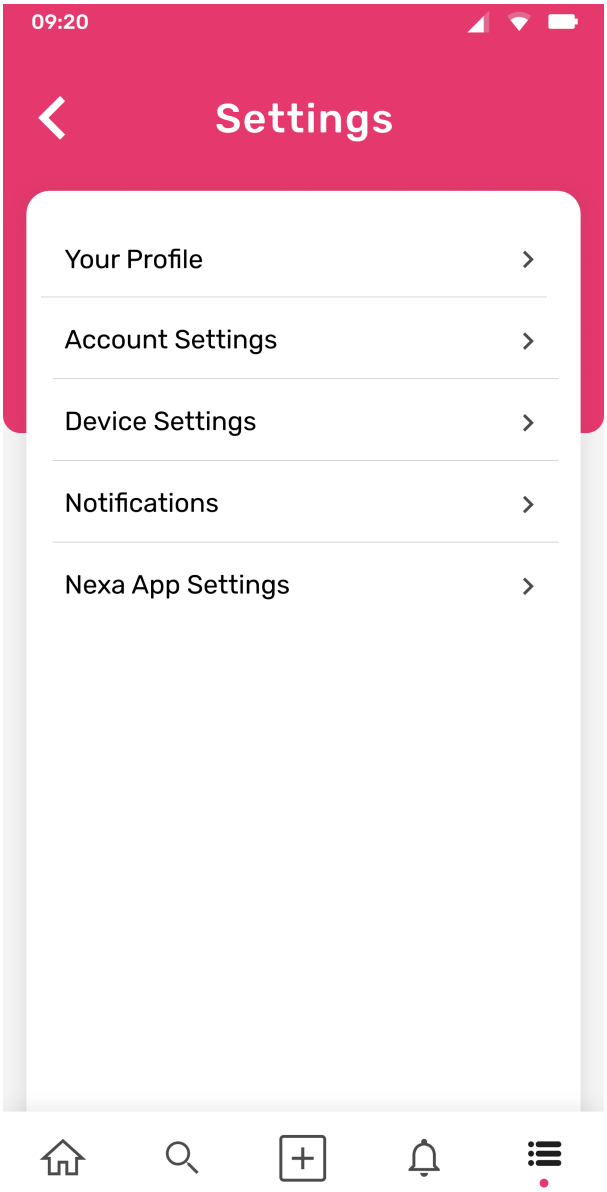
| | | | |
|---|---|---|---|
| Current Privacy Practices | Bad Experience: Read Privacy Policy | [Bad Experience] | The specific bad designs / elements distract the participant from finding / reading the privacy policy |
| Current Privacy Practices | Change Privacy Settings | [name]/No | The name of the settings changed / Haven't changed the privacy settings before |
| Current Privacy Practices | Reasons: Change / Not Change Privacy Settings | [reason] | The participant's reasons for changing / not changing privacy settings |
| Current Privacy Practices | Good Experience: Change Privacy Settings | [Good Experience] | The specific good designs / elements help the participant from finding / changing the privacy settings |
| Current Privacy Practices | Bad Experience: Change Privacy Settings | [Bad Experience] | The specific bad designs / elements distract the participant from finding / changing the privacy settings |
| Current Privacy Practices | Access Data Collected | [name]/No | The name of the data accessed / Haven't access the data collected before |
| Current Privacy Practices | Reasons: Access / Not Access Data Collected | [reason] | The participant's reasons for accessing / not accessing data collected |
| Current Privacy Practices | Good Experience: Access Data Collected | [Good Experience] | The specific good designs / elements help the participant from accessing data collected |
| Current Privacy Practices | Bad Experience: Access Data Collected | [Bad Experience] | The specific bad designs / elements distract the participant from accessing data collected |
| Current Privacy Practices | Delete Data Collected | [name]/No | The type of the data deleted / Haven't deleted the data collected before |
| Current Privacy Practices | Reasons: Delete / Not Delete Data Collected | [reason] | The participant's reasons for deleting / not deleting data collected |
| Current Privacy Practices | Good Experience: Delete Data Collected | [Good Experience] | The specific good designs / elements help the participant from deleting data collected |
| Current Privacy Practices | Bad Experience: Delete Data Collected | [Bad Experience] | The specific bad designs / elements distract the participant from deleting data collected |
| Prototype: VUI | Attitude: General | Satisfied / Not Satisfied | Whether the interaction experience of Nexa meets the expection of the participant or not |
| Prototype: VUI | Reason: Satified / Not Satified | [reason] | The reasons why the participant believes that Nexa provides a good / bad response |
| Prototype: VUI | Attitude: Adequate Information | Adequate / Not Adequate | Whether Nexa provides adequate / not adequate information overally |
| Prototype: VUI | Reason: Adequate / Not Adequate | [reason] | The reasons why the participant believes that Nexa provides adequate / not adequate information |
| Prototype: VUI | Preferred Modality: Privacy Notice | VUI / GUI | Whether the participant would like to receive privacy notice in VUI or GUI |
| Prototype: VUI | Preferred Modality: Privacy Notice | [reason] | The reasons for choosing VUI / GUI as a preferred modality for receiving privacy notice |
| Prototype: VUI | Preferred Modality: Privacy Choice | VUI / GUI | Whether the participant would like to do privacy choices in VUI or GUI |
| Prototype: VUI | Preferred Modality: Privacy Choice | [reason] | The reasons for choosing VUI / GUI as a preferred modality for doing privacy choice |
| Prototype: VUI | Suggestion: Nexa Reply | [suggestion] | The suggestions which the participants gave to improve Nexa's reply |

| | | | |
|---|---|---|---|
| Prototype: VUI | Willingness to Use | Yes/No | Whether the participant would like to ask priacy-related questions to Nexa in daily life, if they can use Nexa in daily life |
| Prototype: VUI | Reason: Willing / Not Willing | [reason] | The reasons for willing / not willing to ask priacy-related questions to Nexa in daily life |
| Prototype: VUI | Scenario: Nexa Privacy Questions | [name] | In what scenarios does the participant would like to ask privacy-related questions to Nexa in daily life |
| Prototype: VUI | Suggestion: Improve VUI Interaction | [suggestion] | The suggestions which the participants gave to improve the overall experience of interacting with the VUI |
| Prototype: GUI | Good Experience: [task name] / General | [Good Experience] | The elements / design patterns help the participant to find the targets for the tasks |
| Prototype: GUI | Bad Experience: [task name] / General | [Bad Experience] | The elements / design patterns distract the participant from finding the targets for the tasks |
| Prototype: GUI | Suggestion: Improve GUI Interaction | [suggestion] | The suggestions which the participants gave to improve the overall experience of interacting with the GUI |
| VUI & GUI Comparsion | Preferred Modality: [task name] / General | VUI / GUI | Whether the participant prefer to use the VUI / GUI for [task name] / in general |
| VUI & GUI Comparsion | Reason: Prefer VUI for [task name] / in General | [reason] | The reasons for preferring VUI over the GUI for [task name] / in general |
| VUI & GUI Comparsion | Reason: Prefer GUI for [task name] / in General | [reason] | The reasons for preferring GUI over the VUI for [task name] / in general |
| VUI & GUI Comparsion | Trust Privacy Notice | VUI / GUI | Whether the participant would trust VUI / GUI more in receiving privacy notice |
| VUI & GUI Comparsion | Reason: Trust VUI / GUI for Privacy Notice | [reason] | The reason for trusting VUI / GUI more in receiving privacy notices |
| VUI & GUI Comparsion | Trust Privacy Choice | VUI / GUI | Whether the participant would trust VUI / GUI more in doing privacy choices |
| VUI & GUI Comparsion | Reason: Trust VUI / GUI for Privacy Choice | [reason] | The reason for trusting VUI / GUI more in doing privacy choices |
| VUI & GUI Comparsion | Suggestion: General | [suggestion] | The suggestions for improving the current practices on privacy notices and choices of voice assistants |

## A.5 Nexa : GUI

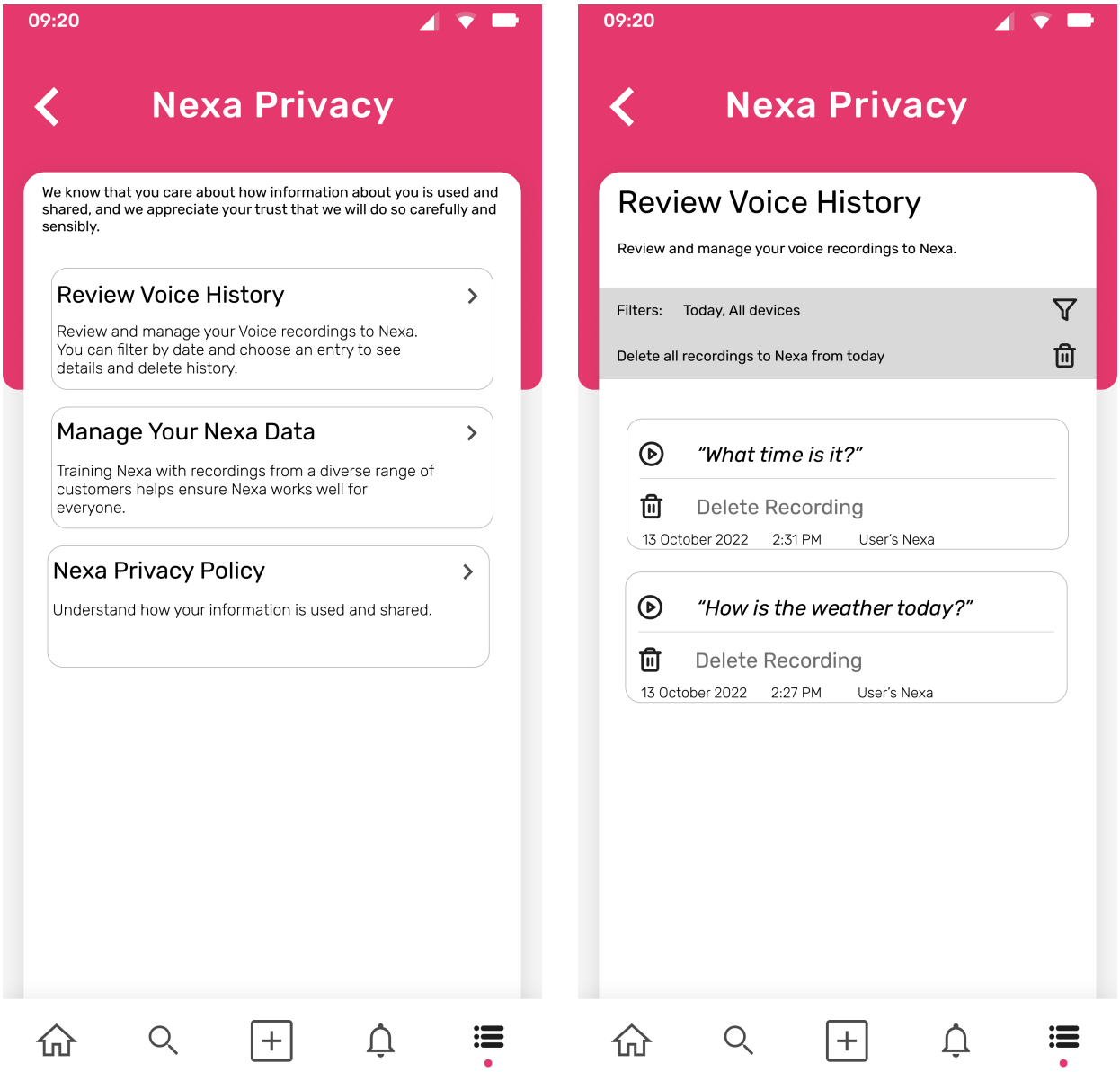



(a) GUI start page




(b) GUI settings page

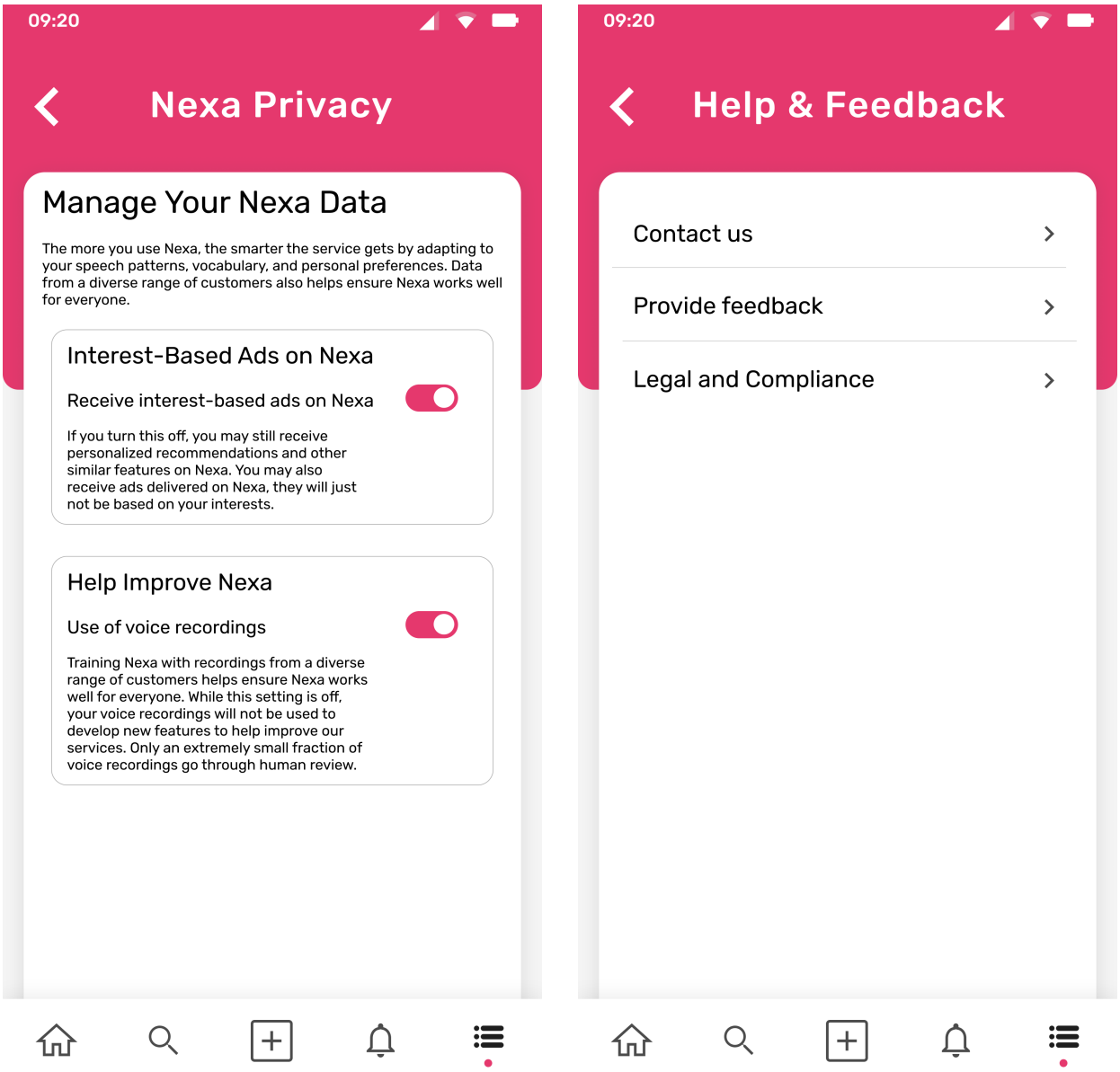
Figure 1: GUI start page and setting page




(a) GUI Nexa Privacy page

(b) GUI Review Voice History page

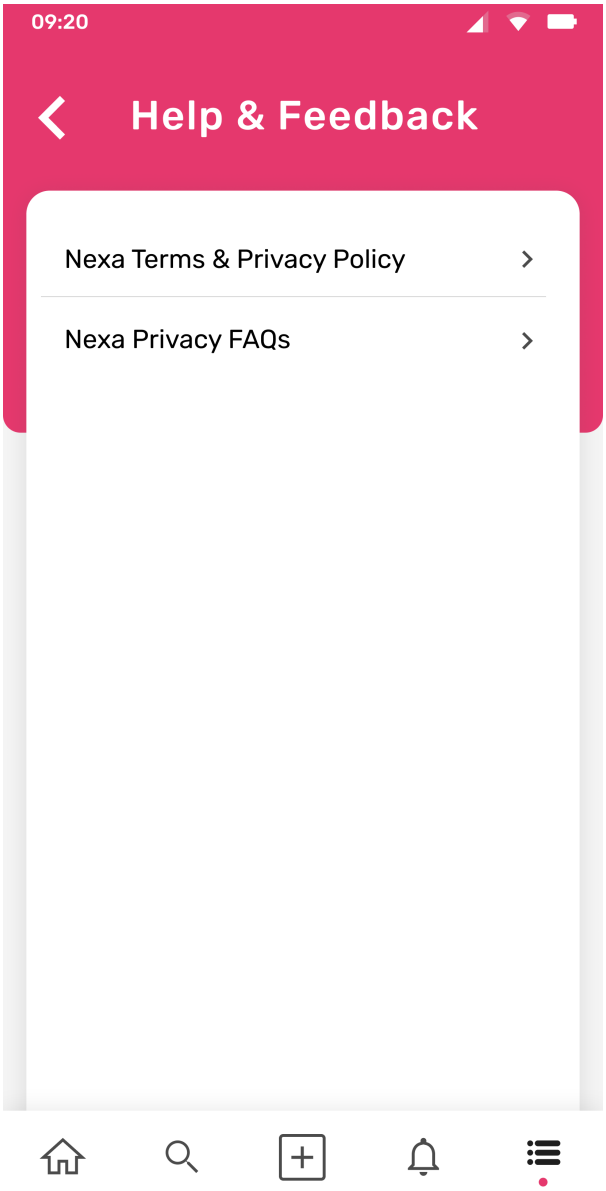
Figure 2: GUI Nexa Privacy page and Review Voice History page



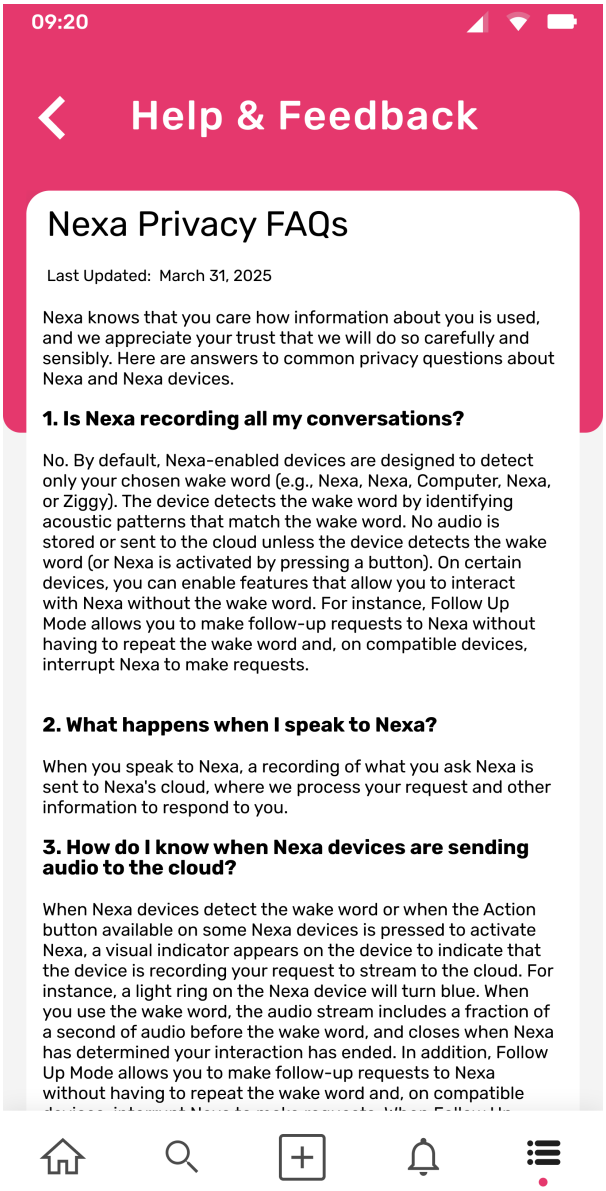

(a) GUI Manage your Nexa Data page

(b) GUI Help page

Figure 3: GUI Manage your Nexa Data page and Help page




(a) GUI Legal and Compliance page




(b) GUI privacy FAQs page

Figure 4: GUI Legal and Compliance page and privacy FAQs page

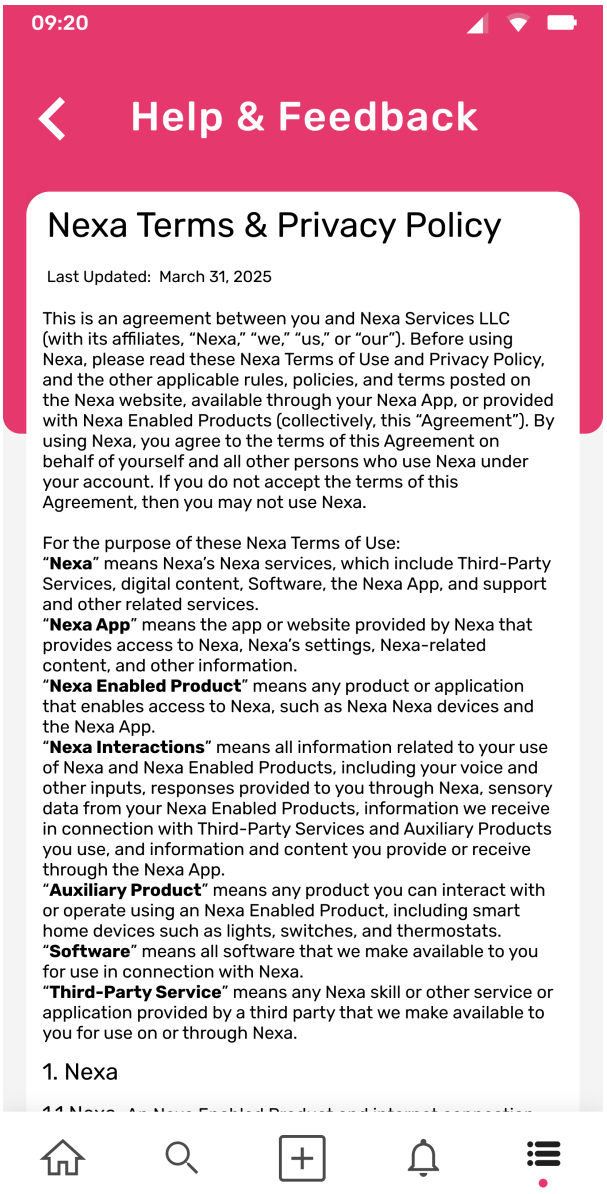



Figure 5: GUI Terms & Privacy Policy page