Music, Search, and IoT: How People (Really) Use Voice Assistants

4 authors, including:

Tawfiq Ammari
University of Michigan
18 PUBLICATIONS **489** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   New fatherhood and social media View project

Music, search and IoT: How people (really) use voice assistants

TAWFIQ AMMARI, University of Michigan School of Information, United States

JOFISH KAYE, Mozilla, United States

JANICE Y. TSAI, Mozilla, United States

FRANK BENTLEY, Verizon Media, United States

Voice has become a widespread and commercially viable interaction mechanism with the introduction of voice assistants (VA) such as Amazon's Alexa, Apple's Siri, Google Assistant, and Microsoft's Cortana. Despite their prevalence, we do not have a detailed understanding of how these technologies are used in domestic spaces. To understand how people use voice assistants, we conducted interviews with 19 users, and analyzed the log files of 82 Amazon Alexa devices, totaling 193,665 commands, and 88 Google Home Devices, totaling 65,499 commands. In our analysis, we identified music, search, and IoT usage as the command categories most used by voice assistant users. We explored how VAs are used in the home, investigated the role of VAs as scaffolding for Internet of Things (IoT) device control, and characterize emergent issues of privacy for voice assistant users. We conclude with implications for the design of voice assistants and for future research studies of VAs.

CCS Concepts: • **Human-centered computing** → **HCI theory, concepts and models**;

Additional Key Words and Phrases: Conversational agents; voice assistants; intelligent assistants; ubicomp; IoT; alexa; google home; home automation.

1 INTRODUCTION

In 1960, J.C.R. Licklider, a computing luminary whose vision laid the groundwork for interactive information systems, posed the question,

> How desirable and how feasible is speech communication between human operators and computing machines?

The question of feasibility for speech communication in human-computer interaction has gone from the realm of science fiction to real life, with voice assistants (VAs) that are commercially-available and widely adopted. There are a variety of assistants across several form factors, ranging from standalone devices (Amazon's Alexa, Google Home), to mobile phone and desktop-based agents (Apple's Siri, Microsoft's Cortana). In fact, a recent Pew poll [43] reports that 45% of Americans use digital assistants, mostly on their smart phones.

Authors' addresses: Tawfiq Ammari, University of Michigan School of Information, Address, City, State, XXX, United States, email; Jofish Kaye, Mozilla, City, United States, email; Janice Y. Tsai, Mozilla, Street, City, State, United States, email; Frank Bentley, Verizon Media, Street, City, State, United States, email.

In spite of high levels of VA adoption, there is a gap in understanding how these technologies are being used in an ongoing basis. Corporations typically do not report on how customers are using their products. Recent HCI panels and workshops (e.g., Kaye et al. [22]) posed questions and introduced research interests in better understanding how VAs are being used and how they can be better designed.

To address this gap, we used multiple methods to triangulate our understanding of the practices of people who had voice assistant devices in their homes. We began by interviewing 19 voice assistant users selected from 132 people recruited from Reddit, focusing on relevant subreddits, like /r/Alexa, and /r/googlehome. Through the interview process, we asked users about their daily use of VAs in order to defamiliarize their use of the technology, thus making their use of VAs more transparent [4]. Defamiliarization allows researchers to "make strange" the assumptions held by technologists about appropriation of technology in domestic spaces, thus allowing us to incorporate "the messiness of everyday life" into our analysis [5].

We found existing research characterizing use of voice assistants at scale was based on self-reported surveys [28, 49], content analysis of user reviews online [49, 50], and interviews [30][1]. It is recognized in data collection literature that self-reporting behaviors could be inaccurate, particularly so when it comes to characterizing one's own behavior over time [63]. In order to address this shortcoming, we triangulated interview and survey responses with data from Amazon Alexa and Google Home history logs. We used Mechanical Turk and Reddit to recruit users who were willing to share the log files from their voice assistant devices and to answer a short survey. We also conducted interviews with a subset of survey respondents.

By combining qualitative data from interviews, and quantitative data from surveys and data logs, our digital traces can be contextualized. This analysis provided us with a macro view of the categories of long-term VA use through log analysis. Much like the "ethnomining" method where digital traces are informed by various different sources of ethnographic data (e.g., interviews), we "extend the social, spacial and temporal scope of research" into daily use of voice assistants [2]. In essence, our qualitative data provided the guidelines to the iterative categorization of commands in user history logs. In doing so, we answer the call of McMillan et al. [32] to use a "combinative method" when studying technology in order to understand its use, not in isolation, but "in interaction."

In our analysis, we found that the three most frequently used command categories are: (1) Music; (2) Hands-free search; and (3) IoT control (e.g., controlling smart lights using voice commands). Our respondents integrated voice assistants into their daily domestic routines, especially when doing so allowed them to carry out their routines more efficiently. For example, some users created voice-activated routines lowering the lights and playing soothing music to help them sleep. Most of our respondents (~70%) knew about the existence of VA history logs, and 10% studied the logs to understand their interactions with the VAs. Most respondents could not articulate specific privacy concerns. When they did articulate specific privacy concerns, respondents discussed: (1) not being sure when the voice assistant is "listening"; and (2) worries about sharing their information with undisclosed third parties when using VAs. While some users relied on privacy controls like muting their VAs to seek more privacy, others were more resigned to their privacy concerns and tended to trust companies operating this emerging technology.

---

[1]See this "What the Amazon Echo is actually used for" info-graphic (https://www.voicebot.ai/2016/10/11/statista-amazon-echo-actually-used/) as an example.

2 RELEVANT WORK

The literature addressing HCI and speech interaction spans a variety of bodies of work and various disciplines such as artificial intelligence, computer science, natural language processing, human-robot interaction, and social psychology. We identified three domains in which we anchor this research: voice assistants, domestic ubicomp and privacy. We use each of these domains to articulate a research question.

2.0.1 A note on language. Some of the names used to describe speech systems or interaction mechanisms include "computers as social actors" (CASA), "conversational agents" (CA), "intelligent personal assistants" (IPAs), "intelligent agents" (IA), and "spoken dialog systems." Furthermore, the various Alexa products, the Google Home products, and similar devices, are sometimes referred to as 'smart speakers'. This diversity of terms (sometimes used interchangeably in a single paper or article) reflects the breadth of research into voice assistants. As a solution, we use the term "voice assistant" in this paper to refer to all speech-driven interaction systems, including Alexa, Google Assistant, Cortana and Siri, which appears to be emerging as an industry standard term.

2.1 Voice Assistants

From a social psychology perspective described in the "computers as social actors (CASA)" literature, Nass et al. [41] conducted experiments to illustrate the intrinsic nature of speech as a driver of social interaction, even with computers. Beyond social interaction, an infrastructure must exist so that the speech can be processed, interpreted, with relevant responses produced for the user. This infrastructure can be described as the spoken dialogue system [34], which is broken out into its technical components (speech technologies, language-processing, dialogue modeling, and processing ability) that enable a user to interact with a complex computer application in a natural way. A key characteristic of interaction is the ability to engage in dialogue with a human user. Conversational agents (CA) or intelligent personal assistants (IPAs) are built on top of spoken dialogue systems. They are often endowed with "humanlike" behavior [61] with a significant focus on the capacity "to carry out tasks." The conversational or intelligent nature is also contingent on the ability for the system to interact in a way that illustrates that it is able to understand context and have a connected interaction across a sequence of conversational turns.

Different researchers have proposed various design principles for VAs. For example, Schechtman and Horowitz [57] focused on tasks, conversation, and relationships, and observed that task completion can impact user satisfaction [24]. Similarly, Porcheron et al. [48] examined the use of VAs in situ to better understand how people make sense of their conversations with a VA (Siri, in this case). They noted that users interact with VAs as though they were "humanlike conversational agents," and suggested that users will build a relationship with their VA [30, 57]. In a study of Amazon reviews of Alexa, Purington et al. [50] quote a reviewer: "Alexa is my new BFF." The VA was not only used for accessing information or entertainment, but also as a companion for the user. In addition, VAs allowed users to collaborate when using VAs to search for information [48]. Guha et al.[19] suggest three factors for successful continued interactions with VAs: (1) contextual assistance such as using the location of the user; (2) content and updates based on user interests; and (3) personalization, using context (defined as tasks, ongoing interests, and routines) to provide suggestions [19].

While Guha et al. did not focus on the use of VAs at home, Porcheron et al. [47] analyzed how families interacted with voice assistants in situ. Based on earlier work by Reeves and Brown [52], the authors analyzed how "the Echo is made 'at home' and 'embedded' into various activities of home life." Similarly, Rode et al. [53] argue that, "in domestic ubicomp, programming becomes a household responsibility, [much like] loading the dishwasher and taking out the trash." The need

for "programming" arises from the fact that new domestic technologies are not used in isolation from the "complex domestic environments in which they are situated" [53]. Tolmie et al. [58] argue that "when digital resources enter the home they cannot just be positioned in any way within the household and its routines," but also depend on collaborative action by the users engaged in using VAs [48].

The studies discussed above were not long term analyses of VA use. For example, Porcheron et al. [47] conducted a month-long study analyzing the use of Amazon Alexa logs and an audio recorder. An earlier study by Porcheron et al. [48] has focused on the use of voice assistants in short term interactions in public spaces. Other VA studies like that by Purington [50] and Luger [30] relied on interviews and surveys which did not make use of VA logs in their analyses. A longer term understanding of how people are using their voice assistants in everyday life is still lacking. Therefore, our first research question is

RQ1: What are the daily uses of VAs?

2.2 Introducing IoT devices to the home

Voice commands have been part of visions of smart homes in movies, television and literature for at least the last fifty years [6]. In the 1990s, voice commands features as part of the Intelligent Room Project at MIT [11] where commands could be issued verbally to different parts of the room (e.g., lights). House_n [21] and later AwareHome [23] were two laboratory studies of domestic ubiquitous computing. Early critiques of projects like the MIT Intelligent Room Project focused on the affordances that current technology can provide for users (e.g., [9]). However, since these technologies were not widely deployed at the time, studying them in the wild would have been a challenging undertaking. However, one survey shows that 1.1 million IoT systems were installed in US homes throughout 2012[2]. Some of these newly installed IoT systems include smart lights (e.g., Phillips Hue Lights[3]), thermostats (e.g., Nest[4]), stereo systems (e.g., Sonos[5]) and cameras (e.g., Nest Cam[6]).

While earlier studies focused on analyzing the use of IoTs in laboratory settings [21, 23], Mennicken and Huang [35] build on Bell and Kaye's [6] view that studying ubiquitous systems should focus on the experiences of the users, rather than the creation of efficiencies in domestic spaces like the kitchen. They study user experiences in relation to domestic routines [14], other actors in the home and the technology affordances [35, 46], thus defamiliarizing the system's use [4]. The authors found that users install IoT systems when they found such systems to be convenient - a finding that echoes that of Brush et al. [12]. Other users wanted to live in modern homes which "should have the highly advanced technological infrastructure, even when their ideas about such infrastructure were vague [35]." One concrete reason given by users for employing IoT systems was in the area of savings (e.g., using a smart thermostat to reduce heating fuel consumption).

A Convenient system is one that "fits, speeds up, or improves" family routines [35]. Mennicken and Huang found that users employed IoT devices to "hack" the home and make their routines flow better. Mennicken and Huang define drivers as those who push the hacking process at home, but, as opposed to Poole et al. [46], they find that other members of the household tended to be passive users, rather than helpers in hacking the home. One of the reasons for this role breakdown might be related to the lack of a central operating system to control the multitude of IoT systems [17].

---

[2]https://www.abiresearch.com/press/15-million-home-automation-systems-installed-in-th/
[3]https://www2.meethue.com/en-us
[4]https://nest.com/thermostats/nest-learning-thermostat/overview/
[5]https://www.sonos.com/en-us/home
[6]https://nest.com/cameras/nest-cam-indoor/overview/

New IoT platforms tend to be heterogeneous, thus raising the cost of interacting with them and connecting the different IoT devices [36, 54, 58, 64]. While some users interact with a multitude of apps to control different IoT devices, others install gateways or hubs that allow them to communicate with and connect different IoT devices [58, 59]. These hubs allow users to create macros to control IoT devices and use information across the different IoT platforms [35]. Setting up these IoT devices requires significant technical work by the users. In fact, some hubs assume coding knowledge to set up macros for using different IoT devices.

Tolmie et al. [58] refer to the labor associated with setting up and maintaining IoT devices at home as "digital plumbing." With the addition of more IoT devices, these technologies need to be incorporated into domestic routines [15]. This incorporation into the family daily routine can be complex as family members discover new ways to implement their routines with each added IoT device and iterate to include more IoT devices in their smart homes [35].

While there have been studies analyzing the use of IoT devices in smart homes, we lack an understanding of the ways VAs are used in relation to other IoT devices in smart homes. Therefore, we ask the following research question,

RQ2: How do users incorporate voice assistants into their IoT domestic setup?

2.3 Privacy

Cloud-connected or "always on" systems introduce new challenges for maintaining users' privacy. Data, and its collection, use, and sharing are often invisible. It is very difficult to design and deploy privacy-sensitive ubicomp systems [20]. Since the current legal framework around privacy is based on a notice and consent model that "cannot hope" to meet the challenges posed by ubiquitous computing systems [29], a new system of communication for privacy preferences and consent are needed. Other methods of presenting terms and conditions for mobile and ubiquitous technology was proposed by Morrison et al. [39] where the use of the system would be interrupted with "visual representations of collected data" as opposed to long descriptions of such data.

Earlier work suggests that there are privacy concerns specific to the use of voice assistants. Diao et al. [16] discuss security problems that show how voice assistant components are potential security threats. Moorthy and Vu [38] discuss privacy issues that arise from using voice assistants in public such as being overheard. Indeed, privacy preferences are often nuanced and context dependent. Naeini et al. [40] found people were uncomfortable with IoT-based data collected in their homes and with data shared with 3rd parties. Oulasvirta et al. [44] studied the long term effects of surveillance using different modalities (e.g., video camera and smart phones) in one's domestic environment and found that users changed their behaviors to reduce privacy violations (e.g., not walking naked even in the privacy of their own home). The reason for these changes in behavior can be explained by a privacy concept heavily relied on in the HCI literature, namely boundary regulation [45]. Boundary regulation is a process of socio-technical negotiation between individuals, groups of people who might be affected by technology use (e.g., family and friends), and technology designers [45]. In the case of VA and IoT devices, the negotiation is between primary users, usually the ones who setup and configure emerging technologies around the home, and secondary users like other family members, friends, or roommates [26].

Relatedly, the theory of privacy as contextual inquiry stipulates that privacy needs change according to the social context [42]. Klasnja et al. [25] describe how privacy concerns depended on the type of information collected, the context of collection, and the value derived from collecting the information. For example, audio recording in professional settings, especially when intimate information is shared (e.g., recording in a psychologist's office) are deemed unacceptable. On the other hand, data that allows users to track their exercise are deemed more acceptable.

Given that earlier work describes a number of privacy concerns specific to ubiquitous technology, we ask

> **RQ3a: What privacy concerns do users of voice assistants have when incorporating the new technology in their daily interactions?**

Consumers want their data to be used for purposes that can provide them with actual value. Once this kind of information (data collection, use, and sharing) is made available and users are able to have control, they often decide to allow personal information to be shared. One solution to consider for the future is the use of personalized privacy assistants which could make privacy choices on behalf of the user based on previous privacy preferences [27]. Designing for transparency, awareness, and control is important, but can be difficult to accomplish. Lau et al. [26] argue that design of VAs "did not align" with the privacy needs of users. Users thought that privacy controls like the history logs and mute button were cumbersome and difficult to conceptualize.

As users incorporate voice assistants into their daily routines, we ask

> **RQ3b: What privacy controls did VA users employ to mitigate their privacy concerns? How did they perceive VA privacy controls?**

3 METHOD

To understand how people use voice assistants, we conducted interviews with 19 participants to explore how voice assistant users made sense of these new technologies. We then collected Amazon Alexa and Google Home "histories," automatically generated logs of commands, to analyze patterns of use, ultimately analyzing 82 logs totaling 193,665 commands for Amazon Alexa, and 88 logs totaling 65,499 commands for Google Home. These logs were categorized into several main command categories. Our surveys and data collection mechanisms were approved by our organizations' review processes.

3.1 Interviews

3.1.1 Recruitment. We recruited interviewees via Reddit. After contacting Reddit moderators to introduce our project, we asked if we could post our recruitment messages to their respective boards. We posted a message on several subreddits that have users interested in home networking, voice assistants, and IoT devices in general (e.g., r/Alexa, r/googlehome, r/HomeAutomation). The recruiting message contained a link to an online screening survey in SurveyMonkey, soliciting people over the age of 18 based in the United States. We asked for information about VA technologies used and collected demographic information. We interviewed 19 out of a total of 132 respondents to the survey. See Table 1 for an overview of our interviewees.

3.1.2 The Interviews. Interviews were conducted between June 20th and June 24th, 2017. The median length of the interviews was 39.5 minutes, with a standard deviation of 11.3 minutes. Interviewees were provided with a $100[7] Amazon.com gift certificate as a token of appreciation for their participation. Respondents recruited via Reddit may be more technically capable than the average user. However, since we are studying the use patterns of a relatively new technology, the viewpoint of highly motivated and technically savvy users are useful in understanding how users might implement the use of VAs in general. We started each interview by asking the respondents about the devices they identified in the survey. We moved to focus on their use of the Internet, including their thoughts and concerns around privacy. We then asked about how respondents used their voice assistant(s) as well as any IoT devices or ubicomp technologies they used on a daily basis domestically.

---

[7] While this value might be high for some academic studies, it is in line with the values paid for research subjects in industry.

Alias	M/F	Age	State	Kids?	VA devices used	IoT Devices
Molly	F	28	IL	No	(AA, 1), (S,1)	None
Brad	M	63	TX	Yes	(AA, 4)	Smart switches,smart lights, Harmony hub
Boris	M	26	NY	No	(AA, 2)	Smart lights
Bob	M	30	MA	No	(AA, 2)	Smart lights, Harmony Hub
Chuck	M	30	IL	Yes	(AA, 1), (S,1)	Smart switches, smart lights, smart lights
Mona	F	25	CA	No	(AA, 2)	Smart lights, Nest
Hari	M	25	WA	No	(AA, 1), (GH, 1)	Smart lights, Nest
Harriet	F	36	CO	No	(AA, 4), (GH, 1)	Smart lights, Nest, Smart humidity sensor
John	M	24	FL	No	(GH, 1)	Smart lights
Duke	M	19	VA	No	(AA, 2)	Smart lights
Daniel	M	40	PA	Yes	(S, 1)	Smart lights, Nest, smart lock
Kyle	M	23	CA	No	(AA, 1)	Smart lights
Susan	F	26	WA	No	(AA, 1)	Smart lights, Nest
Jose	M	26	FL	No	(AA,1), (GH, 1)	Smart switches
Gavin	M	33	SC	Yes	(GH, 4)	Smart lights, Nest, smart smoke alarm, smart switch, smart lock, Harmony Hub
Monique	F	43	AZ	No	(GH, 1)	Smart switches
Mark	M	29	GA	No	(GH, 4)	Smart lights
Timothy	M	29	GA	No	(GH, 1)	Smart lights
George	M	43	IN	Yes	(AA, 1), (GH, 2)	Smart lights

Table 1. Interviewee details. AA: Amazon Alexa, GH:Google Home, S:Siri

3.1.3 Analyzing the interviews. The interviews were transcribed and the transcripts coded using NVivo, a qualitative data analysis package[8]. The interviews were analyzed using an inductive process in which the first author conducted multiple passes, discussing the emerging codes after each pass with co-authors.

The themes included discussions of how parents used hands free search as well as music commands. They also included descriptions of how interviewees set-up their IoT environment and used VAs in conjunction with it. We also asked users to describe their interactions with other members of the family when using VAs. Finally, users discussed privacy concerns they might have when using VAs. In the results section, we expand on the themes shown in Table 2 to show how the interviewees conceived of their use of voice assistants in their everyday lives.

---

[8]http://www.qsrinternational.com/nvivo/nvivo-products

Theme	# interviews theme is discussed	# times theme is discussed
Search	17	42
Music	16	39
Timers	14	20
Internet of Things	16	51
Smart home and IoT hubs	8	15
Macros and programming	8	13
Family interactions	15	120
Privacy	19	24

Table 2. Key codes for interview analysis

3.2 Surveys

Previous research by Bentley et al. [8] has shown that using samples of participants from Mechanical Turk can be reliable in understanding technology use when compared to large-scale professional market research surveys or the analysis of usage logs held by large corporations. Given the time and expense of collecting thousands of logs, we believe this method provides a dataset that allows us to analyze the use of these devices in the wild.

Similar methods have been used by in earlier work to analyze the use of mobile devices, specifically cell phone use. Bentley and Chen [7] use survey data along with data from user smart phones to analyze their interactions with their social networks, while Battestini et al. [3] analyze similar questions through collecting all the text messages sent and received by the study participants. In both studies, the authors noted that log collection allow researchers to collect data without the potential disadvantage of missing entries (e.g., when respondents forget to enter data in diary studies).

We used MTurk and Reddit to recruit users who wanted to receive $5 in return for filling out a short survey and sharing the logs of their voice assistant usage. Questions on the survey included how long they had owned the device and where the device was located in the home. The survey concluded by capturing basic demographic information including the composition of their household.

The users were asked to answer a question about their geographic location. Since the voice assistant logs store timestamps in Unix epochs[9], these data were used to localize the timestamps from each of the user logs. We also allowed respondents to provide some free-text responses discussing their experiences with VAs. The survey took an average of six minutes to complete.

We summarize survey responses in Table 3 below. Of the Amazon Alexa user respondents, 26% identified as female, while 47% of the Google Home sample identified as female. Respondents covered an age range of 18-64 years. The respondents to the Google Home sample were more likely to be the sole member of their household, and were drawn from a smaller number of US states (27 vs. 37). It is not clear if the demographic differences between these two samples is indicative of patterns in the users of the two products.

---

[9]https://www.unixtimestamp.com/index.php

	Female * (%)	Sole Household(%)	No. of States **	Age range	No. of Logs
Amazon Alexa	26	17	37	18-56	82
Google Home	47	31	27	18-64	88

Table 3. Summary of survey results. * All other respondents selected male as a gender. None of our respondents chose gender non-conforming or other. ** This represents the number of states where respondents live in the United States

3.3 History logs: Dataset

We used Amazon Mechanical Turk and Reddit to recruit participants to provide us with full device usage logs from 82 Amazon Alexa users and 88 Google Home owners. We performed the data collection in a manner similar to the phone book data collection study by Bentley and Chen [7]. Participants were provided with detailed instructions on how to access their Amazon Alexa or Google account history on the respective web pages for each product[10]. Participants were given the opportunity to remove any entries that they did not feel comfortable sharing with the research team.

3.3.1 Amazon Alexa Logs. We collected a total of 193,665 commands on Amazon Alexa between May 4^{th}2015 and August 2nd 2017, a period of 851 days. On average, the datasets for our 82 Amazon Alexa users span 210 days. On the days when they used their VA, Alexa users issued, on average, 18.2 commands per day with a median of 9.0 commands per day.

3.3.2 Google Home Logs. In total, we collected a total of 65,499 commands on Google Home collected between September 21st 2016 and July 10th 2017, a period of 293 days. On average, the datasets for each of our 88 Google Home users spans 110 days. On days when they used their VA, Google Home users issued, on average, 23.2 commands per day with a median of 10.0 commands per day.

Google Home users issued 5 more commands on active days than did their Amazon Alexa counterparts. We do not have a hypothesis as to why this is.

3.3.3 Defining command categories. In our analysis, we used the Python Pandas library. Pandas is an "open source library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language."[11] We read the logs into a Pandas data frame, with each row representing a command. The columns for each command included:

- the command text: This is the text used in our categorization. An example would be "Alexa, play music."
- time stamp for command: We used timestamps to determine the density of certain commands throughout the day
- name of the device: This column identifies the name of the device the user directed the command to. We have removed this column from our analysis in order to maintain the privacy of our respondents. Many of the devices contained some identifying information (e.g., name of the user or names of family members).

We began searching the dataset based on the themes that arose through the qualitative exploratory analysis of the interview data. We then found all the commands related to each of these categories (and sub-categories). In order to check the main commands in each category, we found the highest frequency terms and applied TF-IDF[12] to the commands to find the terms with the highest scores.

---

[10]https://alexa.amazon.com, https://myactivity.google.com
[11]https://pandas.pydata.org/
[12]Term frequency-inverse document frequency is a score used to find the most important words in a corpus of documents.

TF-IDF determines the relative frequency of words in a document as compared to the inverse proportion of that word in the complete corpus. This would increase the score of words that occur more rarely throughout the corpus as opposed to "common words such as articles and prepositions [56] cited in [51]". In our case, these would be wake words like Alexa since they are repeated at a high rate throughout the log data.

If any of the words with the highest TF-IDF scores were unrelated to the category, they would be added to the list of comments in another category and removed from the category currently being analyzed. We then checked a number of commands picked randomly to make sure that the commands are indeed part of the category. Defining the command categories was an iterative process. Each iteration allowed us to hone the command category further through analyzing other related commands.

For example, we describe how we analyzed the command category, Music. We started by looking for commands containing the seed words "play, pause, stop, resume, restart, shuffle." As we sampled commands from different logs allowed us to build on the command criteria if the commands are deemed to have a similar user intent. After the first iteration, we found that there are other commands that, while using some of the terms in the regular expression above, do not relate to playing music. For example, we found that some of the users were "**play**ing" a skill called Jeopardy. Others played the news. One of the log entries we had not anticipated here was "Text not available. Click to **play** recording." This is the Alexa log entry signifying that Alexa is unable to parse the audio data. After finding these exceptions and a few others, we added another regular expression to exclude them from the music criteria. Finally, we analyzed the highest-frequency words as well as the terms with the highest term frequency-inverse document frequency (TF-IDF) score. If any of the most popular terms (high TF-IDF score) were not related to the category in question, then we would incorporate that information into the regular expression. The next iteration allows us to have a more precise categorization of commands presented in the command logs. After a number of iterations, we created a category resembling a group of commands, in this case, music-related commands. The complete example with associated code is presented in Appendix A.1.

4 RESULTS

Below, we analyze the findings of our study. We start by describing some of the main uses of Amazon Alexa and Google Home. We then discuss the effects of incorporating a VA on the IoT environment at home. We also discuss privacy concerns users have when using voice assistants.

Figure 1 shows the breakdown of the command categories for Amazon Alexa and Google Home. We can see that Search, Music and IoT commands are the three most frequently used command categories for both Amazon Alexa and Google Home. In order to further show examples of the commands within command categories, Table 4 shows the most frequent words, and most popular terms for command categories used most heavily in the Amazon Alexa files. Table 5 shows the most frequent words, and high score TF-IDF terms for command categories used most heavily in the Google Home command logs.

Out of the 193,665 Amazon Alexa commands, we found that 51,491 commands consisted only of wake words like "Amazon", "Alexa", "Echo," and "Computer". Ford and Palmer [18] have previously reported that Alexa devices will sometimes spontaneously wake without intentional invocation from a user, and this result may confirm that finding. We did not find an equivalent command category in the Google Home logs. For the analysis below, we omitted these Alexa wake-only commands, which comprise 26.5% of the total, so that we could compare Amazon Alexa and Google Home usage directly.
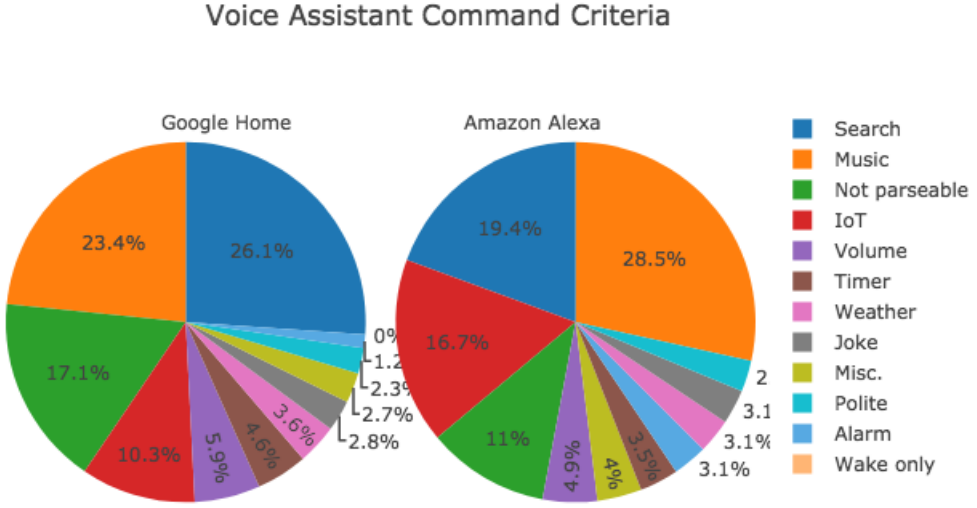
Voice Assistant Command Criteria

Fig. 1. Breakdown of command categories on Google Home and Amazon Alexa

4.1 Music, media and volume

Based on the log analysis, playing music was the most common use of Amazon Alexa (at 28.5%) and the second most used command category for Google Home (at 26.1%). Users played music based on genre (e.g., classical music), album (e.g., "The Fame" by Lady Gaga), or artist (e.g., The Beatles). Voice assistant users also employed Spotify, Pandora, and other music streaming services when listening to music. Duke notes,

> I use Pandora pretty heavily, so, once in a while I'll just have Alexa put on whatever Pandora station I have or want to listen to.

However, another user noted that using Spotify on Amazon Alexa has some limitations, namely that it does not "play my own host play-lists on Spotify." The importance of music as a VA command category is best exemplified by one respondent who wrote,

> it's almost sad to think that we only use it for music.

Kyle noted that he uses blue-tooth speakers connected to Amazon Alexa to play music in different parts of the house. Other respondents suggested that they might use Alexa for sounds related to specific routines. For example, one Alexa user noted, "I mainly use my Alexa at night right now for sleep sounds." The use of Alexa to access music also determined its physical location at home. For example, Gavin noted that

> My wife loves music and is a music teacher, so she loves music, listens to music all the time. She also loves to cook and bake, so it made the most sense [to place Alexa] in the *kitchen*.

Figure 2a shows the heatmap for the music and search on Amazon Alexa aggregated over the 24-hour time line. We present the weight of the specific command category as a portion of all other commands throughout that period of time,

$$Command\ Weight = \sum_{i=0}^{T} Music\ Commands / \sum_{i=0}^{T} All\ Commands$$
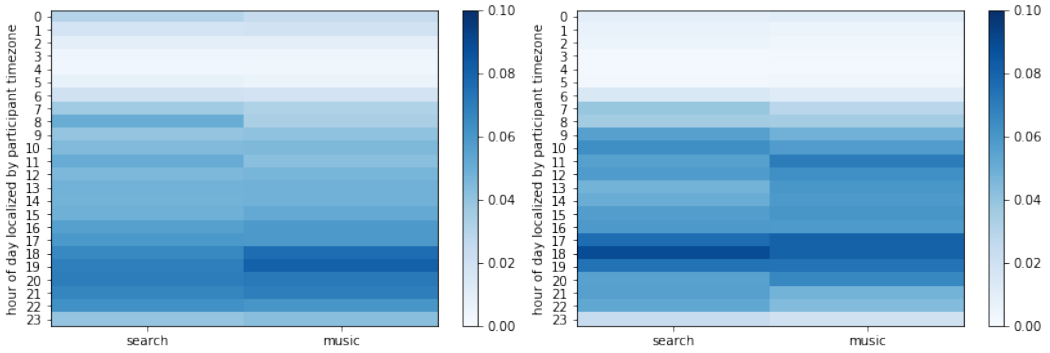
(a) Heat map for search and music commands aggregated over hour of day for Amazon Alexa. The value represents the weight music/search commands over the total commands in the same period of time. Music commands tend to be most dense between 6 and 8 pm.

(b) Heat map for search and music commands aggregated over hour of day for Google Home. The value represents the weight music/search commands over the total commands in the same period of time. Music commands tend to be most dense between 4:30 and 6:30 pm.

Fig. 2. Search and music commands hourly usage for Amazon Alexa and Google Home

where T is fixed at one hour intervals

For Amazon Alexa, the music command was used most heavily between 6 and 10 pm, while peaking between 6 and 8 pm. Figure 2b shows a the equivalent heatmap for Google Home music and search commands over the 24-hour time line. Similar to the Amazon Alexa heatmap, we find that music was used most heavily between 6 pm and 9 pm. This might arise because users are listening to music while preparing meals at the end of the workday.

Because it was so common, we pulled volume out as a separate category from music. Around 4.9% of Amazon Alexa interactions and 5.9% of Google Home commands were volume related. Curiously, we found that the ratio of "volume up" to "volume down" commands for Alexa was 37% and 30% for Google Home commands, suggesting that both Alexa's and Google Home's default volume may be set too high.

Interviewees did not limit their voice assistant use to music. Some interviewees indicated that they used their voice assistant to access other media. For example, Jose noted how he used Google Home, along with Google Chromecast[13] to operate his Netflix account. Brad discussed how he used Amazon, along with Harmony Hub[14] to control his entertainment center:

> there was a lot of remotes involved [with the entertainment center]. It's the kind of thing where someone comes over to your house and they can't figure out how to run the system...now you can say alright CNN is a channel...so in the future if I say 'Turn on CNN', it'll turn to that channel. That's somewhat useful, but mainly I use it for to turn on/off, for muting and unmuting, and for pausing and resuming.

Other interviewees were introduced to smart home devices as they integrated smart speakers (e.g., Sonos) to be used with their voice assistants. Chuck notes that "the fact that [Alexa] could be auxiliary plugged into a Sonos Play 5 was appealing to me...They are always expanding their skillset [sic] and there's a big open source community around building integrations to Alexa." Using

---

[13]https://store.google.com/product/chromecast_2015
[14]https://www.logitech.com/en-us/product/harmony-hub

VAs led users to use smart home devices (e.g., Harmony Hub) for the purpose of interaction with their media environment. Brad describes the layout of VAs in the house thusly

> We have a Harmony Hub for our downstairs entertainment system. We have two Echo Dots. One in the basement. One in the kitchen, which is basically the family room. That one's connected to a Sonos Play 5, so we use that for all of our music streaming and entertaining the kiddo. We have another Sonos Play 1 upstairs.

In deciding where to place VAs, users consider where they listen to music or consume media throughout their daily routines. We further analyze the use of smarthome devices along with VAs in Section 4.4.

4.2 Search

Search or informational queries was the most prevalent use of Google Home (at 26%) and second most prevalent use for Amazon Alexa (at 19.4%). The frequency of search command use was highest for both Amazon Alexa and Google Home was between 5 and 7 pm followed by the time between 8 am and noon.

As Tables 4 and 5 show, one of the most popular terms was "song" for both Amazon Alexa and Google Home. Users used the search command to ask questions about music they listened to, specifically the name of a song they are listening to, or the name of the artist singing a particular song, etc. One respondent notes,

> Oh yeah, a couple times I've used it to identify a song because it's able to do like, Alexa what's that song that goes like and then you just sing a couple verses.

Some respondents emphasized the use of the search feature when interacting with family and friends. For example, Hari commented that "sometimes I have friends around and I could ask random questions, like trivia questions, or like some facts." For Duke, using Alexa to search online served as way to brag to his friends. Other search commands focused on sports scores. They also used the feature to search for trivia ("How many people live in Shelbyville, KY?") or check stock market value ("What's Facebook stock at?"). The heatmap for weights of search commands can be seen in Figure 2a for Amazon Alexa and Figure 2b for Google Home.

Other users noted that search featured in their daily routines. For example, Brad describes how part of the reason they decided to place the VA in the kitchen was that his wife

> uses [VA] a lot for cooking. She uses it for converting measurements, you know how many teaspoons are in a cup... She gets pretty good responses when she asks for substitute ingredients, like if she runs out of something.

The use of search in the process of cooking might be one reason for the higher density of search commands on both Amazon Alexa and Google Home (see Figure 2). Other search queries included asking about movie show times, time when a store closes, when is a person's birthday or a the date of a specific event.

However, queries did not always go as expected: for example, one respondent noted "She can't hear me when the music is playing too loud." But that was not the only problem respondents identified with using search with Alexa.

Brad compared the search feature for Aamazon Alexa and Google Home

> "The main knock on the Echo is that it's not as good as the Google Home for web searches and whatever, but...if I want to Google something, I'll use a computer."

This view was echoed by Hari and Jose, both of whom compared Amazon Alexa and Google Home search. That might explain why Google Home is used more when employing the search command category.

4.3 Timers, jokes, conversations and more!

We can also see that the use of timer command category in both Google Home and Amazon Alexa logs is between 5 and 7 pm. This corresponds to the time users might be cooking dinner at the end of the workday. For example, Gavin notes that they use timers mostly for cooking purposes. Another user says describes how using VAs for timers is better than using dial timers,

> I use timers when I'm cooking. I will say it is so much more convenient for me to do it verbally than it is to, Oh wait, wheres my phone? Oh wait, where's the little dial timer?

Timers could also be used to set reminders for users. Table 5 shows an example where a user sets a timer to "remind me to make a smoothie at 11 am." Monique, who has ADHD, said that she used timers to stay focused on the task at hand.

> I am very ADHD…okay? When I'm doing things, because it's so easy for me to get sidetracked, I do 15-minute timers. Like, let's say I'm filing or working on a paper or something, because it's like I can do anything for 15 minutes, you know what I mean? And so that's thing one is to help keep me on track. It goes off, I go, 'Okay, I've worked 15 minutes." I can feel justified with taking a break and going back to it.

Molly placed her Amazon Alexa "in the living room on top of the coffee table…because that's where we spend most of the time, and it's right next to the kitchen, so I'm always asking to set up alarms." That location also allowed Molly to place items on the shopping list as they ran out in the kitchen.

However, Harriett noted that adding items to the shopping list from different Echos result in redundant items on the list. She wanted the VA to check items across different lists. On a similar note, George noted that his Echo did not provide support multiple users in the same household;

> I have my Google calendar linked in to the Echo, but it's only my calendar. My wife can't have a separate calendar that she uses. She'd have to just use mine as well.

Most interviewees also noted that they used VAs as Alarms. The terms used in VA logs as presented in Tables 4 and 5 show the Alarm category includes words like "set" as in set the Alarm and "snooze" when snoozing the alarm when triggered.

Users also asked about the temperature on that particular time as well as future forecasts, at times asking for a specific day, for example, "Alexa, is it gonna snow two days from now?"

In addition to these functional uses of VAs, respondents also made use of their VAs to interact with other members of the family (for example, parent with children), or to socialize with visitors. Our logs show that users asked Alexa for jokes, told Alexa to meow, or bark (Table 4). Similarly, Google Home users asked their VAs whether they "have a lover?" or if it can "scratch their backs" while also asking for jokes (Table 5). Similar interactions included asking Alexa to read a bedtime story or asking what Alexa's favorite robot is. Similar questions were also common when friends visited and interacted with the VA. For example, Harriet says

> my friends, usually they just talk to it and see if they can trip her up on something. That's really the main game is just to see what stupid tasks they can do with her, see if they can make her curse.

While Monique's friends also try similar fun uses of the VA, "The ones who do not already have some type of home automation device think it's just really wild, cause I walk into a dark room and then all of a sudden the lights are on and they're like, What!"

4.4 How voice assistants motivated home automation

IoT commands were the third most uttered commands for both VAs. IoT commands constitute around 10% of Google Home commands and 16.7% of the Amazon Alexa commands. Both Amazon

Alexa and Google Home provide some form of home automation integration. One respondent commented that his voice assistant, an Alexa, provided "many integration points it has with home automation products and account linking abilities with other services makes it a very useful product for me." The effect of buying a VA was to motivate owners to use these integration points.

Most of the IoT commands for Amazon Alexa (85%), involved switching lights on and off:"Echo, bedside off." The next command group (about 10%) involved dimming the lights and changing light colors, "Alexa, dim lights to twenty percent." Finally, a smaller minority of the commands involved changing the temperature in different parts of the house. For example, "Set kitchen temperature to seventy-six degrees." Similarly, 85% of IoT commands on Google home also referred to switching smart devices on and off, with 10% changing light colors, dimming lights and changing fan speeds.

First, we identify some of the motivations behind the use of IoT commands through VAs. Brad explains that he was originally looking for a Bluetooth connected speaker for his bedroom so that his wife can listen to music. Being an Amazon Prime member,

> I went to Amazon and was looking at bluetooth speakers. It was when they were introducing the first Echo, and they had the $99 deal. I'm always up for a bargain, and it sounded like it would do what I wanted...anyway, once I got the Echo, I started looking into home control.

Monique said that one reason she started investing in smart home appliances after buying Google Home is that she "felt silly to have a $130 clock radio! But I wanted to minimize my buy-in [to home automation] by installing the cheaper smart switches as opposed to smart lights like the Hue." Monique added more IoT devices with time. Other users considered purchasing a VA only if the VA could provide value to the IoT devices installed in their homes. For example, Daniel has been accruing IoT devices, mainly smart lights, but does not own an Alexa. At this point, he would rather control his devices with their individual apps. After listing the different IoT devices and respective apps, he concluded, "I don't really have all the stuff that [Alexa] can control to make it worthwhile yet." Jose expands on the question of considering IoT purchases and says:

> it could be difficult to get to the light in my bedroom especially at night when going to the bathroom...I thought it would be cool to control it...that's why I got the smart plug [switch]...it was really easy to connect to Google Home. Both my girlfriend and I use it. But not sure how to think about buying more of this technology.

As users installed more IoT devices, the need for more VAs in different parts of the house arose. This incremental process of adding more IoT devices and similarly scaling up with more VAs in different parts of the house was discussed by Gavin

> Just as we started using it more, we recognized it'd be more useful in other places, so we got the one in the living room. As we started getting a lit bit into home automation for voice control and then just as that kept growing, we wanted more in each room

Another use of VAs was vaguely related to saving money and energy. Jose noted that he used Google Home with the Nest thermostat when he visited his brother's house. He plans to buy a Nest thermostat to use it along with Google Home when he "becomes a homeowner." This is reflective of view that adding home automation functionality around VAs adds to the value of user homes.

4.4.1 *IoT integration is not without its problems.* Survey respondents noted that they had faced problems while integrating their home automation devices. When asked if they had any other thoughts on automation, one of the free-text responses notes, "not sure why automation with Hue [lights] is so complicated." Another noted that Alexa "needs more integration with other devices." This view was echoed in interviews as well. Duke's experience with connecting IoT devices from different manufacturers meant that he had to use a smart hub to connect the different devices to a

VA. Duke wished he could "see all connected devices without needing to go through a smarthub." Harriet also noted that she had some problems integrating Alexa with her IoT devices. She explains that "Alexa gets confused because there are two different accounts for the same device. For example, [there might be] two kitchen lamps." The reason for the duplication usually has to do with using multiple apps and/or smart hubs to control and integrate IoT devices at home. Harriet had been using both SmartThings (a Samsung IoT device hub) and Wink (another IoT device hub).

On another note, Brad noted that "it would be great" if the VA could understand the context of the command. For example, Brad was interested in having the VA better interpret his comments with relation to his location in the house at the time of issuing the command to Alexa, "when I'm in the living room and I ask Alexa to shut down the lamp, I want her to shut down the lights in the living room." Similarly, another respondent commented that "'Alexa turn on bedside lamp' could mean a different lamp based on who says it."

4.4.2 Advanced IoT functionality: Macros and routines. Five of the respondents created IoT triggers that can be initiated using Alexa. Throughout the Amazon Alexa command logs, there were only 338 triggers used throughout the Amazon Alexa logs. A trigger is the command used to initiate an IoT Hub macros. For example, a user might program a "Play Xbox" trigger that would turn on the TV, Xbox, and stereo,and dim the living room lights. To create or change macros, the user would have to make updates in the software used to manage their IoT hub.

"I love Alexa, although there's lots of things I wish it could do that it can't (like single commands to play music and trigger home automation functions)" commented one of the survey respondents in the free-text question. Similarly, Brad noted that he would like to be able to give Alexa multiple commands at the same time. He noted that at this point, he does so in a "kludgy way" using the SmartThings hub. But he wants the ability to set up macros using Alexa without the need to have a hub as moderator. In addition, interviewees discussed their view that integration could be expanded to involve not only IoT devices, but also media devices like Plex and the Harmony Hub. John expanded on this idea, explaining

> You know,...I wish I could setup custom voice macros...Right? Because then I could really kind of take it to the next level where I could say, set up a party mode...playing party playlist on your Plex and changing the lights to the party mode pattern and doing this and doing that. It's the difference between having a house that could be remotely controlled versus having a house that's truly automated and intelligent.

As can be seen in Figure 3a and 3b, IoT use increases in the evening and again early in the morning. This is the time when family members return home from work in the evening and when they prepare to leave for work. It is also the time when users would start using commands related to putting lights on/off, starting/stopping fans or changing thermostat settings. The weights in the figure represent the weight of the IoT and Timer command categories over the total commands issued to the VA at the same time of day.

4.5 VAs and privacy

Amazon and Google have sought to provide history logs for their users in order to enhance their experience using the VA. Activity logs also provide some transparency and control around data collection (Alexa History and the Google Activity dashboard). Users are able to view the transcription of audio clips, listen to the audio, see Alexa or Google's response, and delete items. We found that survey respondents reported they were aware (69.5%) of the log history, although only a small percentage reported that they had ever deleted any of their log entries (10.9%). Some of the interviewees noted that they used their logs to review their interactions and make sure that
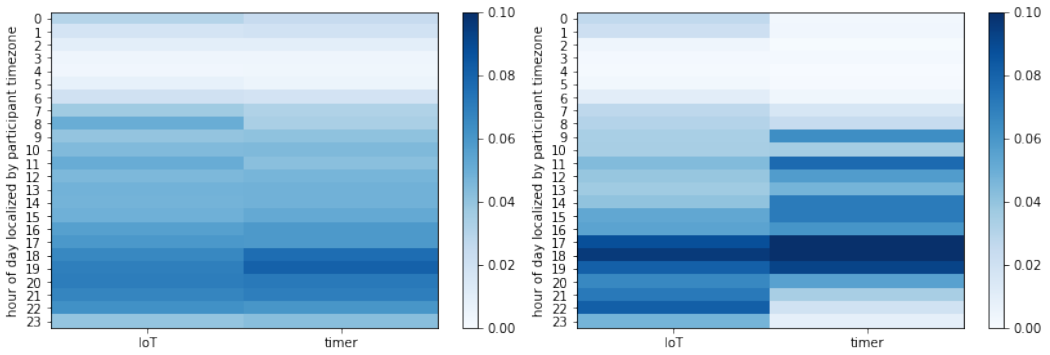
(a) Heat map for IoT and timer commands aggregated over hour of day for Amazon Alexa. The value represents the weight of IoT/timer commands over the total commands in the same period of time. The weight of IoT commands tend to increase around 5 pm, peaking between 9 and 10 pm. Timer commands are issued at higher rates between 6 and 8 pm, presumably for cooking dinner.

(b) Heat map for IoT and timer commands aggregated over hour of day for Google Home. The value represents the weight of IoT/timer commands over the total commands in the same period of time. The weight of IoT commands tend to increase around 6 pm, peaking between 9 and 10 pm. Timer commands are issued at higher rates between 6 and 8 pm, presumably for cooking dinner

Fig. 3. IoT and timer commands hourly usage for Amazon Alexa and Google Home

there were no unexpected interactions. Some users like Monique actually thought the logs helped her better understand her needs. For example, if she

> ask[ed] for a little bit of information about something and then when seeing it in my history go, 'Oh yeah, that was something that interested me. Let me see if there are any books available on that or if there are any movies,' and it inspires me to research further.

However, over one quarter of the survey respondents reported that they did not know that they could delete items in History (26.8%).

Most of our respondents noted that they did not have particularly salient privacy concerns when using VAs. Gavin thought that since "they're waiting for the trigger words, and they can't send any audio before that word is triggered," he does not have privacy concerns specific to using his VA. John, intimating that while he thought there might be some privacy concerns, noted that

> you basically have a microphone that's listening 24/7. It's the same concept of why carrying around a cellphone constantly is the worse possible thing that could ever happen, but it makes life convenient. The primary reason I chose the Google Home over [Amazon Alexa] is because I buy pretty heavily into the Google Eco-system.

In other words, since he had already bought-in to the Google platform, using another product under said platform mitigated John's privacy concerns.

When they did express privacy concerns, those mainly could be broken down to three main themes: (1) Amazon Alexa/Google Home listening to conversations even when not triggered with a wake word ; (2) conversational records that are processed and stored on external machines; and (3) access to private information by third party services (e.g., Amazon Alexa weather skill).

4.5.1 Is she always listening? A survey respondent noted that Alexa sometimes "randomly lights up or is 'listening' when I haven't spoken to her." Their concern is specific to the control/or lack thereof over when their VA is on/off. Molly, who had an Alexa in the living room, similarly noted that there are topics she would prefer not to discuss around Alexa, like family finances and other issues of a personal nature. Another respondent, Mona explains "I prefer to mute her all the time unless we're actually using her for something." Mona expanded on that point by noting that "basically, if we're having sex we mute Alexa. Just in case [because] sometimes she'll start blinking..." without a wake word. Mona was referring to a device the couple had in their bedroom. She followed this comment by saying that "[my boyfriend] thinks I'm paranoid" for muting the VA when not in use. This disparity between different home members when contemplating privacy settings have been echoed by other respondents. For example, Brad's wife was worried that having so many microphones across the house would inevitably result in some privacy invasion. Harriet and Brad both said that that are heavy users of Alexa. They both have more than one VA in different parts of the house. However, Brad and Harriet were both criticized by family members for having too many VAs around their homes. Harriet said that her

> in-laws are mortified that someone could hack in and see what I'm doing, but what are they going to learn? They're going to hear me talking to my husband about mundane stuff like hummus recipes and stuff, so I don't care.

4.5.2 VA logs. Other users had more specific privacy concerns. One such concern centered on the availability of records for their interactions with VAs and the location where these records were stored. For example, one survey respondent said he was, "honestly creeped out that [Alexa] stores so much information I was completely unaware of on a website that's easy to hack." This comment was a reaction to the fact that the respondent did not know of the existence of the Alexa History log before he was introduced to it in our study.

4.5.3 Access to data by third party apps. John noted that he was concerned about how VAs "reach out to...third party services" when for example asking about the weather. He is critical of the fact that he knows very little about what information is sent to third party services and how these data are stored and protected. He followed this comment by saying that he would rather have "locally hosted" systems where he can be in better control of his data. Similarly, one of our survey respondents suggested that he wished for a "an open-source locally-hosted alternative" VA for domestic use.

5 DISCUSSION

In this section, we reflect on our findings and how they relate to earlier work in this space. First, we discuss daily VA use by analyzing the main command categories for Amazon Alexa and Google Home users (RQ1). Next, we discuss how users incorporated VAs into their IoT domestic setup (RQ2). Finally, we analyze our findings with regards to privacy concerns and measures users take to protect their privacy when using VAs (RQ3).

5.1 RQ1: What are the daily uses of VAs?

With our analysis of Alexa and Google Home History and Activity data, we have a more concrete and accurate understanding of how people are using their VAs (especially compared to self-reported usage). We found that the three main uses for both Google Home and Amazon Alexa are (1) music, (2) hands free search, and (3) IoT control, primarily turning lights on and off. We also introduce some of the less frequently used command categories.

Following, we address our findings related to each of the categories. We also address some of the less prevalent command categories and how they were used around the house.

5.1.1 Music. VA provided users with the ability to play music. This music could be related to a particular genre (e.g, classical music), written by a particular composer or artist (e.g., the Beatles) or a particular song (e.g. "Just Dance" by Lady Gaga). Users also played music from music streaming streaming services like Pandora and Spotify. Playing music could also be related to users' daily routines. For example, one of our interviewees suggested that he used his VA to play music that he sleeps to. Another indicated that part of the reason his family decided to place a VA in the kitchen is that his wife, a musician, liked to listen to music whilst cooking. This finding echoes earlier results in Volokhin and Agichtein [62] which show that contextual music recommendations depend on the activity the user is undertaking at home. For example, the music one plays when cooking might be different from that they play when they wanted to sleep, clean the house or play with the children.

5.1.2 Hands free search. Related to the Music category, the search category showed that users asked about music they were listening to: who was singing, when was the song written, and so on. Hands free search also provided users with affordances to conduct hands-free online search throughout their daily routines. Some users searched through recipes whilst cooking, reducing the need to touch devices while working in the kitchen. Other users asked about trivia while hosting friends and family. These different uses affected how users considered where they would place VAs around the house.

The search feature also provided a conversation topic between the owners, other family members and their visitors. For example, users engaged in collaborative search when engaged in trivia or other discussions. This finding echoes results from Porcheron et al. [48] stating that the use of VA "has the effect of democratizing the device use by allowing any member to engage without invitation, and to intervene or collaborate with the unfolding device interaction." Another form of social interactions involving VAs included users who noted that they "brag" to friends and family about the VA, which at times led to the visitors considering to purchase their own VA.

5.1.3 Other uses of VAs. Social interactions with visitors afforded by VAs also extended to having more conversational interactions with VAs including asking for jokes. This echoes findings by Purington et al. [50] where users indicated that users had a personal relationship to their VAs. Users still wanted a more naturally conversational technology but indicated that group conversational experimentation was an important part of their experience. IoT integration with VAs also provided opportunity to discuss the new technology with others. We'll discuss more about IoT integration in the following section.

5.2 RQ2: How do users incorporate voice assistants into their IoT domestic setup?

IoT integration commands represented the third most used command categories in both Google Home and Amazon Alexa VA logs. Both VAs provided users with a chance to extract more value from other technologies in their homes through providing a scaffolding for the management of IoT devices.

Brad started thinking of IoT devices he could add to his home once he set up Alexa and Echo devices. While users may buy and install different IoT devices, the real value they gain out of the use of a VA is the connection between different IoT devices. If one installs a number of different "Things" around the home, the ability to communicate with them without having to access multiple apps is of value. For example, the highest frequency words and words with highest TF-IDF scores show that VAs have been used to control IoT devices in different parts of the house (kitchen, bedroom, living room etc.) mostly to turn lights on and off. While some users found such use convenient,

this idea was a vague one related to making the home a "modern" or "smart home." This echoes the findings of Mennicken and Huang as they studied the introduction of IoT devices in domestic spaces outside of laboratory settings [35].

The way users measured the value of VAs and IoT devices around the home changed in relation to: (1) home ownership and (2) daily routines that could be automated. Respondents noted that they would be more willing to install more IoT devices and more VAs if they owned the house since they thought making their domicile smarter added to its value. Much like respondents in Mennicken and Huang [35], our respondents wanted to identify daily routines that could be made easier, while maintaining a low price-range, when using their VAs in addition to IoT devices. With each iteration, users who found the integration of VAs with IoT devices at home thought that they might want more VAs in different parts of the house to control even more IoT devices. Much like users who iterated their IoT installations in Mennicken and Huang [35], we also found that VA users iterated using more VAs and integrating them with more IoT devices as they made sense of the capacities of both.

However, our respondents still indicated that their use of VAs along with IoTs was not without its problems. Users indicated that their VAs lack contextualization in two main ways: (1) spatiotemporal contextualization; and (2) dynamic instruction contextualization, or macros. Below, we expand on each of these contextualization issues.

5.2.1 Spatial and temporal contextualization. Spatial contextualization refers to the capacity of the VA to recognize where the user is physically at any particular point in time. If a user wants to control an IoT device in the living room while in the living room, then the VA should understand that the user is attempting to control the IoT device in the living room, unless otherwise specified by the user. Similarly, Guha et al. in their design recommendations argued for the importance of geographical contextualization of the data used by the VA [19]. In their case, this contextualization required the use of the GPS coordinates of the user.

Rong et al. [55] tackled a similar problem of temporal contextualization setting up calendar appointments. In their system, the main problem was to allow the VA to make sense of a command such as "remind me to get milk this afternoon." Note that the command here is not specific, but relational. The user is assuming that the VA can contextualize her command in the same way that a human would.

In a similar vein, we suggest the design of a spatial and temporal contextualization for user commands, especially when relayed to IoT devices around the home. This can be done by providing an easy way for the user to dynamically "map" their house, where each VA is available in that map and how it relates to the location of IoT devices. For example, when a user sets one VA in the living room and one in the bedroom, the user could dynamically allocate IoT devices to be controlled by default through the VA. If the user is in the bedroom and wants to change the fan setting, she should not have to specify that the command is referencing the fan in the bedroom, the VA should provide the spatial context. This is especially interesting given that our respondents were installing multiple VAs to control devices in different rooms of the house.

5.2.2 Dynamic instructions. When using VAs to control domestic IoT devices, users indicated that they wanted to dynamically control IoT actions via what two of the users termed "macros." Macros would allow users to control a number of different IoTs in relation to a specific activity. For example, if the user is leaving the house, she might want to turn off lights in the house, close the garage door, and reduce the temperature on the thermostat.

At this point, the only way for users to create macros is by programming them through IoT hubs. As we saw from our results, a small proportion of the commands were trigger specific macros for IoT devices. Instead of having to create macros in gateways and then trigger them using VAs, users

should be able to create macros dynamically as they use their IoT devices. For example, we can envision having a separate wake word that establishes delimiters for the beginning and ending of dynamic macros.

This finding is echoed in work by Mennicken and Huang [35] as they suggest that new IoT systems should support "hackers and the hacking process." As with the users interviewed in their study, the capacity to hack the home, and program VAs as they control more IoT devices, was a major motivation for users as they considered buying new IoT devices, and then in turn more VAs to control IoT devices in different parts of the house.

When introducing new IoT devices like smart thermostats or smart lights, these technologies are not programmed in isolation from other technologies in the home. If the VA provides users with flexible tools to "program" [53] their new devices, it will allow users to more easily engage in digital plumbing of their smart homes [58].

However, current VA designs still have to face a major disadvantage, namely, the lack of universal protocols for different IoT devices [36, 37, 64]. New VA designs can provide better affordances by providing user and geographical contextualization and embedding dynamic programming.

5.3 RQ3a: What privacy concerns do users of voice assistants have when incorporating the new technology in their daily interactions?

Most of the respondents did not articulate a coherent view of any privacy concerns they might have when using VAs. For example Harriet told us she had no privacy concerns, and while John intimated some consternation because of a continually working microphone at home, he is already invested in the Google platform, and explained that therefore adding another device linked to the same platform (Google Home) would not be such a privacy threat to him.

However, other members of the household/family members did have privacy concerns, as expressed by Hariette's in-laws and Brad's wife, especially when there were multiple VAs in different parts of the house. As secondary users of technology introduced to the domestic environment by Brad and Hariette, they had less control over its introduction into the home environment [26]. Even when more than one user can be considered a primary user, as with Mona and her boyfriend, they might have divergent privacy concerns. Indeed, Mona's boyfriend thought that she was paranoid for wanting to mute the VA in their bedroom. These divergences represent a privacy boundary management problem [45]. As VAs are introduced into environments with multiple users who might have different privacy needs while sharing the same physical space, designers could introduce ways to provide users with granular control mechanisms when using VAs in different parts of the house. For example, the VA in the bedroom could be muted automatically after 9 pm until the morning alarm.

Our respondents made it clear that they did not know what information was shared with third party services, or how the data was shared. For example, when using a weather Alexa skill, the users do not have a clear understanding of the data shared with third party weather apps. Following the recommendations of Morrison et al. [39], the use of the VA can be interrupted with a voice message to the user to explain what data is being shared when using third party skills.

5.4 RQ3b: What privacy controls did VA users employ to mitigate any privacy concerns? How did they perceive VA privacy controls?

Our research reinforced our admittedly pre-existing assumption that VA developers need to provide usable and prominent information about how consumers can have control over their data. Some of the survey respondents did not know that the history log existed for their VA, let alone that they could access the log and delete earlier commands and queries. As our results show, only a fraction of those who knew of the existence of the logs edited them for privacy concerns. This

finding echos findings by Lua et al. [26] indicating that while users might know of the logs, they might find accessing and editing them too cumbersome.

One concern that users did explain clearly referred to not knowing whether their VA is listening when they did not want it to listen. For example, Molly physically unplugged the VA when discussing financial issues because she did not trust that Alexa would not be listening if it were muted. Mona made a similar statement talking about VAs in the bedroom. Recent work by Ford and Palmer [18] shows that indeed, when Alexa is muted, it does not record audio and send it to the Amazon Service for processing. However, they found that when not muted, Alexa sometimes does interact with the Amazon service, even when a wake word was not used. It might be important to provide better cues showing that the VA is actually muted. For example, when muted, the VA could display a significantly different color/icon in order for users to be sure that the VA is indeed muted. In addition, the logs could show users when their VA was muted, which might get users to trust their VA in operating in a more predictable way. Further, new designs might provide some cues that show when the VA is interacting with the cloud service.

Some of the respondents who did have privacy concerns were most worried about the fact that their speech is being processed remotely. VA producers can ameliorate the users' concerns by providing detailed information about when and with whom these data will be shared [60].

Another change that VA producers might enact is on-device processing. If speech processing is done locally, there would be no need to send the data outside of the user's network to be processed using cloud services. Users could be advised of the technical limitations of on-device processing. Users may then choose to accept said limitations, or rely on cloud-processing of their utterances.

6 LIMITATIONS AND FUTURE WORK

In this work, we provided an exploratory study of the use of voice assistants in day-to-day activities. As with any other study, our study has its limitations. While the interviews provided a qualitative insight into the use of VAs on a daily basis, they have more limitations when compared to diary entries by users when data is still fresh in the users' memories. They are also less contextualized than in-home interviews at the site of VA use where the researchers can collect more information about the environment in which the VA is used along with other technologies at home. While our recruitment from communities on Reddit allowed us to better understand how early adopters appropriated the technology, future work could focus on recruitment from more varied pools of users.

Future work can focus on the use of technology in relation to family routines. For example, earlier literature studied how parents help their children learn to use VAs [28] and engage in conversation repair mechanisms [13, 33]. An important future study would report on how parents, whose responsibilities include managing children's use of and engagement with and management of technology [1, 10, 31], engage with their children as they have increasing access to VAs?

While we articulated the broad command categories of VA use, future work could focus on the effects of current VA uses on future use patterns. Another area to investigate is the adoption of VAs by different user profiles. For example, can the current use of VAs predict future use of VAs by the user? Does the use of Harmony hubs, which have some IoT characteristics, result in the increased use of IoT devices? As the use of voice assistants like Siri and Google assistant (usually on cell phones) increase [43], how is the use of these technologies affecting the way users think of VAs at home? How does it affect the way they decide whether to adopt VAs at home or not? How does it affect their privacy concerns when using VAs?

7 CONCLUSIONS

As voice assistant (VA) use becomes more widespread, we need a better understanding of daily use of this technology. Drawing on 19 interviews, surveys and the logs from 88 Google Home users and 82 Amazon Alexa users, we provide an exploratory study of the daily uses of voice assistants. We found that the three most frequently used command categories were (1) Music; (2) Search and (3) Information of Things (IoT) control commands. We describe how the incorporation of VAs at home affected the way users thought of incorporating IoT devices and vice versa. We also described how users thought about integrating VAs with IoTs. Finally, we analyze privacy concerns around the use of VAs at home, specifically, knowing when VAs are recording and the opaqueness of cloud-based services used by VAs.

REFERENCES

[1] Tawfiq Ammari, Priya Kumar, Cliff Lampe, and Sarita Schoenebeck. 2015. Managing Children's Online Identities: How Parents Decide What to Disclose About Their Children Online. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1895–1904. https://doi.org/10.1145/2702123.2702325

[2] Ken Anderson, Dawn Nafus, Tye Rattenbury, and Ryan Aipperspach. 2009. Numbers have qualities too: Experiences with ethno-mining. In *Ethnographic praxis in industry conference proceedings*, Vol. 2009. Wiley Online Library, 123–140.

[3] Agathe Battestini, Vidya Setlur, and Timothy Sohn. 2010. A Large Scale Study of Text-messaging Use. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10)*. ACM, New York, NY, USA, 229–238. https://doi.org/10.1145/1851600.1851638

[4] Genevieve Bell, Mark Blythe, and Phoebe Sengers. 2005. Making by Making Strange: Defamiliarization and the Design of Domestic Technologies. *ACM Trans. Comput.-Hum. Interact.* 12, 2 (June 2005), 149–173. https://doi.org/10.1145/1067860.1067862

[5] Genevieve Bell and Paul Dourish. 2007. Yesterday's Tomorrows: Notes on Ubiquitous Computing's Dominant Vision. *Personal Ubiquitous Comput.* 11, 2 (Jan. 2007), 133–143. https://doi.org/10.1007/s00779-006-0071-x

[6] Genevieve Bell and Joseph Kaye. 2002. Designing Technology for Domestic Spaces: A Kitchen Manifesto. *Gastronomica* 2, 2 (2002), 46–62. https://doi.org/10.1525/gfc.2002.2.2.46

[7] Frank R. Bentley and Ying-Yu Chen. 2015. The Composition and Use of Modern Mobile Phonebooks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2749–2758. https://doi.org/10.1145/2702123.2702182

[8] Frank R. Bentley, Nediyana Daskalova, and Brooke White. 2017. Comparing the Reliability of Amazon Mechanical Turk and Survey Monkey to Traditional Market Research Surveys. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 1092–1099. https://doi.org/10.1145/3027063.3053335

[9] Anne-jorunn Berg. 1995. A Gendered Socio-Technical Construction: The Smart House. In *Information technology and society: a reader*, Nick Heap (Ed.). Sage, London.

[10] Lindsay Blackwell, Emma Gardiner, and Sarita Schoenebeck. 2016. Managing Expectations: Technology Tensions Among Parents and Teens. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 1390–1401. https://doi.org/10.1145/2818048.2819928

[11] Rodney A. Brooks. 1997. The intelligent room project. In *Cognitive Technology, 1997. Humanizing the Information Age. Proceedings., Second International Conference on*. IEEE, 271–278.

[12] A.J. Bernheim Brush, Bongshin Lee, Ratul Mahajan, Sharad Agarwal, Stefan Saroiu, and Colin Dixon. 2011. Home Automation in the Wild: Challenges and Opportunities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 2115–2124. https://doi.org/10.1145/1978942.1979249

[13] Yi Cheng, Kate Yen, Yeqi Chen, Sijin Chen, and Alexis Hiniker. 2018. Why Doesn't It Work?: Voice-driven Interfaces and Young Children's Communication Repair Strategies. In *Proceedings of the 17th ACM Conference on Interaction Design and Children (IDC '18)*. ACM, New York, NY, USA, 337–348. https://doi.org/10.1145/3202185.3202749

[14] Andy Crabtree and Tom Rodden. 2004. Domestic Routines and Design for the Home. *Comput. Supported Coop. Work* 13, 2 (April 2004), 191–220. https://doi.org/10.1023/B:COSU.0000045712.26840.a4

[15] Andy Crabtree and Peter Tolmie. 2016. A Day in the Life of Things in the Home. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 1738–1750. https://doi.org/10.1145/2818048.2819954

[16] Wenrui Diao, Xiangyu Liu, Zhe Zhou, and Kehuan Zhang. 2014. Your Voice Assistant is Mine: How to Abuse Speakers to Steal Information and Control Your Phone. In *Proceedings of the 4th ACM Workshop on Security and Privacy in Smartphones & Mobile Devices (SPSM '14)*. ACM, New York, NY, USA, 63–74. https://doi.org/10.1145/2666620.2666623

[17] Colin Dixon, Ratul Mahajan, Sharad Agarwal, A. J. Brush, Bongshin Lee, Stefan Saroiu, and Victor Bahl. 2010. The Home Needs an Operating System (and an App Store). In *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks (Hotnets-IX)*. ACM, New York, NY, USA, 18:1–18:6. https://doi.org/10.1145/1868447.1868465

[18] Marcia Ford and William Palmer. 2018. Alexa, are you listening to me? An analysis of Alexa voice service network traffic. *Personal and Ubiquitous Computing* (June 2018), 1–13. https://doi.org/10.1007/s00779-018-1174-x

[19] Ramanathan Guha, Vineet Gupta, Vivek Raghunathan, and Ramakrishnan Srikant. 2015. User Modeling for a Personal Assistant. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15)*. ACM, New York, NY, USA, 275–284. https://doi.org/10.1145/2684822.2685309

[20] Jason I. Hong and James A. Landay. 2004. An Architecture for Privacy-sensitive Ubiquitous Computing. In *Proceedings of the 2Nd International Conference on Mobile Systems, Applications, and Services (MobiSys '04)*. ACM, New York, NY, USA, 177–189. https://doi.org/10.1145/990064.990087

[21] S. S. Intille. 2002. Designing a home of the future. *IEEE Pervasive Computing* 1, 2 (April 2002), 76–82. https://doi.org/10.1109/MPRV.2002.1012340

[22] Joseph 'Jofish' Kaye, Joel Fischer, Jason Hong, Frank R. Bentley, Cosmin Munteanu, Alexis Hiniker, Janice Y. Tsai, and Tawfiq Ammari. 2018. Panel: Voice Assistants, UX Design and Research. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. ACM, New York, NY, USA, panel01:1–panel01:5. https://doi.org/10.1145/3170427.3186323

[23] Julie A. Kientz, Shwetak N. Patel, Brian Jones, Ed Price, Elizabeth D. Mynatt, and Gregory D. Abowd. 2008. The Georgia Tech Aware Home. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems (CHI EA '08)*. ACM, New York, NY, USA, 3675–3680. https://doi.org/10.1145/1358628.1358911

[24] Julia Kiseleva, Kyle Williams, Jiepu Jiang, Ahmed Hassan Awadallah, Aidan C. Crook, Imed Zitouni, and Tasos Anastasakos. 2016. Understanding User Satisfaction with Intelligent Assistants. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval (CHIIR '16)*. ACM, New York, NY, USA, 121–130. https://doi.org/10.1145/2854946.2854961

[25] Predrag Klasnja, Sunny Consolvo, Tanzeem Choudhury, Richard Beckwith, and Jeffrey Hightower. 2009. Exploring Privacy Concerns about Personal Sensing. In *Pervasive Computing (Lecture Notes in Computer Science)*, Hideyuki Tokuda, Michael Beigl, Adrian Friday, A. J. Bernheim Brush, and Yoshito Tobe (Eds.). Springer Berlin Heidelberg, 176–183.

[26] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, Are You Listening?: Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 102.

[27] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, S. A. Zhang, Norman Sadeh, Alessandro Acquisti, and Yuvraj Agarwal. 2016. Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *Symposium on Usable Privacy and Security*. https://www.usenix.org/system/files/conference/soups2016/soups2016-paper-liu.pdf

[28] Silvia Lovato and Anne Marie Piper. 2015. "Siri, is This You?": Understanding Young Children's Interactions with Voice Input Systems. In *Proceedings of the 14th International Conference on Interaction Design and Children (IDC '15)*. ACM, New York, NY, USA, 335–338. https://doi.org/10.1145/2771839.2771910

[29] Ewa Luger and Tom Rodden. 2013. An Informed View on Consent for UbiComp. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. ACM, New York, NY, USA, 529–538. https://doi.org/10.1145/2493432.2493446

[30] Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA": The Gulf Between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5286–5297. https://doi.org/10.1145/2858036.2858288

[31] Melissa Mazmanian and Simone Lanette. 2017. "Okay, One More Episode": An Ethnography of Parenting in the Digital Age. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 2273–2286. https://doi.org/10.1145/2998181.2998218

[32] Donald McMillan, Moira McGregor, and Barry Brown. 2015. From in the wild to in vivo: Video Analysis of Mobile Device Use. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 494–503.

[33] Emily McReynolds, Sarah Hubbard, Timothy Lau, Aditya Saraf, Maya Cakmak, and Franziska Roesner. 2017. Toys That Listen: A Study of Parents, Children, and Internet-Connected Toys. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5197–5207. https://doi.org/10.1145/3025453.3025735

[34] Michael F. McTear. 2002. Spoken Dialogue Technology: Enabling the Conversational User Interface. *ACM Comput. Surv.* 34, 1 (March 2002), 90–169. https://doi.org/10.1145/505282.505285

[35] Sarah Mennicken and Elaine M. Huang. 2012. Hacking the Natural Habitat: An In-the-Wild Study of Smart Homes, Their Development, and the People Who Live in Them. In *Pervasive Computing (Lecture Notes in Computer Science)*. Springer, Berlin, Heidelberg, 143–160. https://doi.org/10.1007/978-3-642-31205-2_10

[36] Mohammad-Mahdi Moazzami, Daisuke Mashima, Ulrich Herberg, Wei-Pen Chen, and Guoliang Xing. 2016. SPOT: A Smartphone-based Control App with a Device-agnostic and Adaptive User-interface for IoT Devices. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, New York, NY, USA, 670–673. https://doi.org/10.1145/2968219.2968345

[37] John Moore, Gerd Kortuem, Andrew Smith, Niaz Chowdhury, Jose Cavero, and Daniel Gooch. 2016. DevOps for the Urban IoT. In *Proceedings of the Second International Conference on IoT in Urban Space (Urb-IoT '16)*. ACM, New York, NY, USA, 78–81. https://doi.org/10.1145/2962735.2962747

[38] Aarthi Easwara Moorthy and Kim-Phuong L. Vu. 2014. Voice Activated Personal Assistant: Acceptability of Use in the Public Space. In *Human Interface and the Management of Information. Information and Knowledge in Applications and Services (Lecture Notes in Computer Science)*. Springer, Cham, 324–334. https://doi.org/10.1007/978-3-319-07863-2_32

[39] Alistair Morrison, Donald McMillan, and Matthew Chalmers. 2014. Improving Consent in Large Scale Mobile HCI Through Personalised Representations of Data. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14)*. ACM, New York, NY, USA, 471–480. https://doi.org/10.1145/2639189.2639239

[40] Pardis Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Cranor, and Norman Sadeh. 2017. Privacy Expectations and Preferences in an IoT World. (2017).

[41] Clifford Nass, Youngme Moon, B. J. Fogg, Byron Reeves, and D. Christopher Dryer. 1995. Can computer personalities be human personalities? *International Journal of Human-Computer Studies* 43, 2 (Aug. 1995), 223–239. https://doi.org/10.1006/ijhc.1995.1042

[42] Helen Nissenbaum. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life.* Stanford University Press. Google-Books-ID: _NN1uGn1Jd8C.

[43] Kenneth Olmstead. 2017. Nearly half of Americans use digital voice assistants, mostly on their smartphones. (Dec. 2017). http://www.pewresearch.org/fact-tank/2017/12/12/nearly-half-of-americans-use-digital-voice-assistants-mostly-on-their-smartphones/

[44] Antti Oulasvirta, Aurora Pihlajamaa, Jukka Perkiö, Debarshi Ray, Taneli Vähäkangas, Tero Hasu, Niklas Vainio, and Petri Myllymäki. 2012. Long-term effects of ubiquitous surveillance in the home. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 41–50. http://dl.acm.org/citation.cfm?id=2370224

[45] Leysia Palen and Paul Dourish. 2003. Unpacking privacy for a networked world. ACM, 129–136. https://doi.org/10.1145/642611.642635

[46] Erika Shehan Poole, Marshini Chetty, Rebecca E. Grinter, and W. Keith Edwards. 2008. More Than Meets the Eye: Transforming the User Experience of Home Network Management. In *Proceedings of the 7th ACM Conference on Designing Interactive Systems (DIS '08)*. ACM, New York, NY, USA, 455–464. https://doi.org/10.1145/1394445.1394494

[47] Martin Porcheron, Joel E. Fischer, Stuart Reeves, and Sarah Sharples. 2018. Voice Interfaces in Everyday Life.. In *Proceedings of the 36th Annual ACM Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA. https://doi.org/10.1145/2702123.2702325

[48] Martin Porcheron, Joel E. Fischer, and Sarah Sharples. 2017. "Do Animals Have Accents?": Talking with Agents in Multi-Party Conversation. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 207–219. https://doi.org/10.1145/2998181.2998298

[49] Alisha Pradhan, Kanika Mehta, and Leah Findlater. 2018. Accessibility Came by Accident: Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 459.

[50] Amanda Purington, Jessie G. Taft, Shruti Sannon, Natalya N. Bazarova, and Samuel Hardman Taylor. 2017. "Alexa is My New BFF": Social Roles, User Satisfaction, and Personification of the Amazon Echo. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 2853–2859. https://doi.org/10.1145/3027063.3053246

[51] Juan Ramos. 2003. Using tf-idf to determine word relevance in document queries. In *Proceedings of the first instructional conference on machine learning*, Vol. 242. 133–142.

[52] Stuart Reeves and Barry Brown. 2016. Embeddedness and Sequentiality in Social Media. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 1052–1064. https://doi.org/10.1145/2818048.2820008

[53] Jennifer A. Rode, Eleanor F. Toye, and Alan F. Blackwell. 2005. The Domestic Economy: A Broader Unit of Analysis for End User Programming. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05)*. ACM,

New York, NY, USA, 1757–1760. https://doi.org/10.1145/1056808.1057015

[54] Benjamin Romano. 2017. Managing the Internet of Things. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education (SIGCSE '17)*. ACM, New York, NY, USA, 777–778. https://doi.org/10.1145/3017680.3022452

[55] Xin Rong, Adam Fourney, Robin N. Brewer, Meredith Ringel Morris, and Paul N. Bennett. 2017. Managing Uncertainty in Time Expressions for Virtual Assistants. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 568–579. https://doi.org/10.1145/3025453.3025674

[56] Gerard Salton and Christopher Buckley. 1988. Term-weighting approaches in automatic text retrieval. *Information processing & management* 24, 5 (1988), 513–523.

[57] Nicole Shechtman and Leonard M. Horowitz. 2003. Media Inequality in Conversation: How People Behave Differently when Interacting with Computers and People. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 281–288. https://doi.org/10.1145/642611.642661

[58] Peter Tolmie, Andy Crabtree, Stefan Egglestone, Jan Humble, Chris Greenhalgh, and Tom Rodden. 2010. Digital plumbing: the mundane work of deploying UbiComp in the home. *Personal and Ubiquitous Computing* 14, 3 (April 2010), 181–196. https://doi.org/10.1007/s00779-009-0260-5

[59] Peter Tolmie, Andy Crabtree, Tom Rodden, Chris Greenhalgh, and Steve Benford. 2007. Making the home network at home: Digital housekeeping. In *ECSCW 2007*. Springer, London, 331–350. https://doi.org/10.1007/978-1-84800-031-5_18

[60] Janice Y. Tsai, Patrick Kelley, Paul Drielsma, Lorrie Faith Cranor, Jason Hong, and Norman Sadeh. 2009. Who's Viewed You?: The Impact of Feedback in a Mobile Location-sharing Application. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 2003–2012. https://doi.org/10.1145/1518701.1519005

[61] Giorgio Vassallo, Giovanni Pilato, Agnese Augello, and Salvatore Gaglio. 2010. Phase Coherence in Conceptual Spaces for Conversational Agents. In *Semantic Computing*. Wiley-Blackwell, 357–371. https://doi.org/10.1002/9780470588222.ch18

[62] Sergey Volokhin and Eugene Agichtein. 2018. Understanding Music Listening Intents During Daily Activities with Implications for Contextual Music Recommendation. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval (CHIIR '18)*. ACM, New York, NY, USA, 313–316. https://doi.org/10.1145/3176349.3176885

[63] Rick Wash, Emilee Rader, and Chris Fennell. 2017. Can People Self-Report Security Accurately?: Agreement Between Self-Report and Behavioral Measures. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2228–2232. https://doi.org/10.1145/3025453.3025911

[64] Thomas Zachariah, Noah Klugman, Bradford Campbell, Joshua Adkins, Neal Jackson, and Prabal Dutta. 2015. The Internet of Things Has a Gateway Problem. In *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications (HotMobile '15)*. ACM, New York, NY, USA, 27–32. https://doi.org/10.1145/2699343.2699344

A APPENDIX A

Group	Highest frequency words	Top TFIDF terms	Example
Not parseable	"Text not available. Click to play recording."	—	—
Music	pause, spotify, pandora, music, skip, next, song, stop, play, alexa	stop, play, skip, shuffle, song, lullaby, music, sing, radio, pause	shuffle songs by dropping young
Search	many, list, song, echo, left, tell, much, time, alexa	echo, time, find, state, white, know, thing, series, score, twenty	alexa how many hours are in a year, what states have the death penalty
IoT	ten, set, kitchen, percent, bedroom, living, room, alexa, light, turn,	bedside, turn, door, kitchen, set, light, lamp, percent, room, bed	echo bedside off, echo turn on kitchen light
Volume	eight, ten, seven, four, three, six, five, turn, alexa, volume,	echo, turn, volume, three, level	alexa turn the volume to six
Conversational	hello, play, okay, thank, morning, shut up, hey, night, good, alexa	good, morning, thank, okay, shut up, series, bedtime, story, hello, robot	tell bedtime story to [name redacted], alexa who is your favorite robot
Timer	twenty , thirty, ten, left, five, much, set, minute, alexa, time	timer, remind, add, restart, many, count, delete	alexa how many timers do i have set, alexa delete timer
Alarm	pm, five, morning , echo, thirty, wake, six, alexa, set, alarm	alarm, snooze, wake, clear, check, silence, current, Tuesday, disable, status	alexa what's the status of my alarms, alexa snooze
Weather	gonna, outside, like, rain, forecast, tomorrow, today, temp, alexa, weather	temperature, rain, weather, snow, sun, seven, from	alexa what's the seven day forecast, alexa is it gonna snow two days from now
Joke	amazon, another, like, know, spell, knock, alexa, tell, us,joke	echo, tell, joke, like, dog, say, meow, alexa, know	alexa tell me a star wars joke, alexa can you take a long walk off a short pier
Miscellaneous	repeat, cancel, ad, turn, echo, say, open, unknown, play, alexa	echo, gonna, never, close, oh, change, alexa, dance, day	echo can i change your name to alexa, dance off, repeat, alexa open slogan machine

Table 4. This table shows Amazon Alexa command categories along with highest frequency words and high-score TF-IDF scores for each category. We also provide a few examples for each category

Group	Highest frequency words	Top TFIDF terms	Example
Not parseable	Null	—	—
Music	sing, pandora, google, skip, pause, next, music, song, stop, play	play, skip, stop, song, sing, pause, music, next, resume	hey google next track
Search	name, song, stock , make, price, tell, many, left, much, time	code, live, work, score, nba, star, game, point, list, song	what's the name of this song, what's Facebook stock at
IoT	100, table, bedroom, set, kitchen, lamp, living, room, light, turn	turn, room, light, kitchen, set, bathroom, lamp, dim, doorbell, bed	turntable to 50%, turn on bedside
Volume	seven, 30, set, six, three, four, five, 50, turn,volume	volume, loud, turn, level, loud	increase volume two level seven, make it louder
Conversational	night, shut, stop, thank, morning, good, hey, okay, google	Okay, google, shut up, thank, good, story, read, hey	okay google nevermind, shut up, read me a bedtime story
Timer	15, 1, 3, cancel, 20, 5, 10, minute, set, time	set, reset, time, remind, setup	remind me to make a smoothie at 11 a.m. today, cancel timer
Alarm	15, 8, turn, 6, cancel, 30, minute, set, alarm	alarm, snooz, next, check, current, silence, current, Tuesday, disable, status	snooze for 20 minutes, set an alarm for 6 a.m. tomorrow
Weather	rain, going, snow, forecast, tomorrow, like, outside, today, weather, temperature	weather, temperature, forecast, rain, snow, snowflake	how's the weather tomorrow, what's the weather outside
Joke	make, think, like, know, knock, spell, say, tell, joke	say, love, scratch, tell, joke, like	do you have a lover, can you scratch my back
Miscellaneous	true, love, cancel, talk, google, tell, like, call, day, repeat	address, restart, repeat, obituary	tell me about the day, what is the address of the nearest starbucks

Table 5. This table shows Google Home command categories along with highest frequency words and high-score TF-IDF scores for each category. We also provide a few examples for each category

A.1 An example of command category iteration

As an example, we describe how we arrived at the definition of the command criterion, Music. This criteria shows when users are playing music, along with the interactions users might have when playing music, like stopping music, shuffle, pause, or moving to the next song. All of these functions were named by our interviewees as they discussed their use of voice assistants. First, we loaded user logs into one data frame using Python Pandas. This allowed us to search through command logs efficiently.

```
#Using regular expressions library in python
import re
#We identified text criteria that would identify music-related commands
music_criteria=
r'rap|fastforward|rewind|ditty|lullaby|play|pause|
song|sing|skip|stop|music|next|pandora|spotify|listen|radio|resume|restart|shuffle'
```

After the first few iterations, we found that there are other commands that, while using some of the terms in the regular expression above, do not relate to playing music. For example, we found that some of the users were "**play**ing" a skill called Jeopardy. Others played the news. One of the log entries we had not anticipated here was "Text not available. Click to **play** recording." This is the Alexa log entry signifying that Alexa is unable to parse the audio data. After finding these exceptions and a few others, we added another regular expression to exclude them from the music criteria. The next iteration allows us to have a more precise categorization of commands presented in the command logs. After a number of interactions, we created a category resembling a group of commands, in this case, music-related commands.

```
#Identifying commands that are not related to music, but appear in original query
not_music_criteria=
r'^(?!Text not available. Click to play recording.|news|jeopardy|stop the alarm).)*$'
```

For both Google Home and Amazon Alexa logs, we created new data frames for each of the new sub-categories. For example, we created the data frame (df_music) for commands that correspond to regular expressions shown above.

In order to determine the residual miscellaneous category (df_miscellaneous), we excluded all the commands categorized in other data frames. We found nine main command categories in addition to the residual category. All these categories are presented in Table 3 for Amazon Alexa and Table 4 in Google Home.

```
#Referring to the data frame as df
#the new miscellaneous data frame, df_miscellaneous will contain all commands not
#picked in any of the nine command categories identified earlier
df_miscellaneous = df[~df.command.isin(all_commands)]
```