

Music, Search, and IoT: How People (Really) Use Voice Assistants 1 2

TAWFIQ AMMARI, University of Michigan School of Information
 JOFISH KAYE and JANICE Y. TSAI, Mozilla
 FRANK BENTLEY, Verizon Media

Voice has become a widespread and commercially viable interaction mechanism with the introduction of voice assistants (VAs), 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 VAs, 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 VA users. We explored how VAs are used in the home, investigated the role of VAs as scaffolding for Internet of Things device control, and characterized emergent issues of privacy for VA users. We conclude with implications for the design of VAs and for future research studies of VAs. 3
4
5
6
7
8
9
10
11
12

CCS Concepts: • **Human-centered computing** → **HCI theory, concepts and models;** 13

Additional Key Words and Phrases: Conversational agents, voice assistants, intelligent assistants, ubicomp, IoT, alexa, google home, home automation 14
15

ACM Reference format: 16

Tawfiq Ammari, Jofish Kaye, Janice Y. Tsai, and Frank Bentley. 2019. Music, Search, and IoT: How People (Really) Use Voice Assistants. *ACM Trans. Comput.-Hum. Interact.* 26, 3, Article 17 (April 2019), 28 pages. <https://doi.org/10.1145/3311956> 17
18
19
20

1 INTRODUCTION 21

In 1960, Licklider, a computing luminary whose vision laid the groundwork for interactive information systems, posed the question, 22
23

How desirable and how feasible is speech communication between human operators and computing machines? 24
25

The question of feasibility for speech communication in human–computer interaction (HCI) 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 26
27
28
29

Q1

Authors’ addresses:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or to publish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

1073-0516/2019/04-ART17 \$15.00

<https://doi.org/10.1145/3311956>

30 desktop-based agents (Apple’s Siri, Microsoft’s Cortana). In fact, a recent Pew poll [43] reports
31 that 45% of Americans use digital assistants, mostly on their smart phones.

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

37 To address this gap, we used multiple methods to triangulate our understanding of the practices
38 of people who had VA devices in their homes. We began by interviewing 19 VA users selected from
39 132 people recruited from Reddit, focusing on relevant subreddits, like /r/Alexa and /r/googlehome.
40 Through the interview process, we asked users about their daily use of VAs in order to defamiliarize
41 their use of the technology, thus making their use of VAs more transparent [4]. Defamiliarization
42 allows researchers to “make strange” the assumptions held by technologists about appropriation
43 of technology in domestic spaces, thus allowing us to incorporate “the messiness of everyday life”
44 into our analysis [5].

45 We found existing research characterizing use of VAs at scale was based on self-reported surveys
46 [28, 49], content analysis of user reviews online [49, 50], and interviews [30].¹ It is recognized in
47 data collection literature that self-reporting behaviors could be inaccurate, particularly so when it
48 comes to characterizing one’s own behavior over time [63]. In order to address this shortcoming,
49 we triangulated interview and survey responses with data from Amazon Alexa and Google Home
50 history logs. We used Mechanical Turk and Reddit to recruit users who were willing to share the
51 log files from their VA devices and to answer a short survey. We also conducted interviews with a
52 subset of survey respondents.

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

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

¹See 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 73

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: VAs, domestic ubicomp, and privacy. We use each of these domains to articulate a research question. 74
75
76
77
78

Q2

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 VAs. As a solution, we use the term “VA” in this article 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. 79
80
81
82
83
84
85
86

2.1 Voice Assistants 87

From a social psychology perspective described in the “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 characteristic of interaction is the ability to engage in dialogue with a human user. CA or 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. 88
89
90
91
92
93
94
95
96
97
98
99

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 CA,” 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]. 100
101
102
103
104
105
106
107
108
109
110
111
112

While Guha et al. did not focus on the use of VAs at home, Porcheron et al. [47] analyzed how families interacted with VAs 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 113
114
115
116
117
118
119

120 digital resources enter the home they cannot just be positioned in any way within the household
 121 and its routines,” but also depend on collaborative action by the users engaged in using VAs [48].

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

129 *RQ1: What are the daily uses of VAs?*

130 2.2 Introducing IoT Devices to the Home

131 Voice commands have been part of visions of smart homes in movies, television, and literature
 132 for at least the last 50 years [6]. In the 1990s, voice commands features as part of the Intelligent
 133 Room Project at MIT [11], where commands could be issued verbally to different parts of the room
 134 (e.g., lights). House_n [21] and later AwareHome [23] were two laboratory studies of domestic
 135 ubiquitous computing. Early critiques of projects like the MIT Intelligent Room Project focused
 136 on the affordances that current technology can provide for users (e.g., [9]). However, since these
 137 technologies were not widely deployed at the time, studying them in the wild would have been a
 138 challenging undertaking. However, one survey shows that 1.1 million IoT systems were installed
 139 in US homes throughout 2012.² Some of these newly installed IoT systems include smart lights
 140 (e.g., Phillips Hue Lights³), thermostats (e.g., Nest⁴), stereo systems (e.g., Sonos⁵), and cameras
 141 (e.g., Nest Cam⁶).

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

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

158 New IoT platforms tend to be heterogeneous, thus raising the cost of interacting with them and
 159 connecting the different IoT devices [36, 54, 58, 64]. While some users interact with a multitude of
 160 apps to control different IoT devices, others install gateways or hubs that allow them to commu-
 161 nicate with and connect different IoT devices [58, 59]. These hubs allow users to create macros to

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/>.

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 VAs. Diao et al. [16] discuss security problems that show how VA components are potential security threats. Moorthy and Vu [38] discuss privacy issues that arise from using VAs 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?

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

216 As users incorporate VAs into their daily routines, we ask

217 *RQ3b: What privacy controls did VA users employ to mitigate their privacy concerns?*

218 *How did they perceive VA privacy controls?*

219 3 METHOD

220 To understand how people use VAs, we conducted interviews with 19 participants to explore how
221 VA users made sense of these new technologies. We then collected Amazon Alexa and Google
222 Home “histories,” automatically generated logs of commands, to analyze patterns of use, ultimately
223 analyzing 82 logs totaling 193,665 commands for Amazon Alexa, and 88 logs totaling 65,499 com-
224 mands for Google Home. These logs were categorized into several main command categories. Our
225 surveys and data collection mechanisms were approved by our organizations’ review processes.

226 3.1 Interviews

227 *3.1.1 Recruitment.* We recruited interviewees via Reddit. After contacting Reddit moderators
228 to introduce our project, we asked if we could post our recruitment messages to their respective
229 boards. We posted a message on several subreddits that have users interested in home networking,
230 VAs, and IoT devices in general (e.g., r/Alexa, r/googlehome, r/HomeAutomation). The recruiting
231 message contained a link to an online screening survey in SurveyMonkey, soliciting people over
232 the age of 18 based in the United States. We asked for information about VA technologies used
233 and collected demographic information. We interviewed 19 out of a total of 132 respondents to the
234 survey. See Table 1 for an overview of our interviewees.

235 *3.1.2 The Interviews.* Interviews were conducted between June 20th and June 24th, 2017. The
236 median length of the interviews was 39.5 minutes, with a standard deviation of 11.3 minutes.
237 Interviewees were provided with a \$100⁷ Amazon.com gift certificate as a token of appreciation
238 for their participation. Respondents recruited via Reddit may be more technically capable than the
239 average user. However, since we are studying the use patterns of a relatively new technology, the
240 viewpoint of highly motivated and technically savvy users are useful in understanding how users
241 might implement the use of VAs in general. We started each interview by asking the respondents
242 about the devices they identified in the survey. We moved to focus on their use of the Internet,
243 including their thoughts and concerns around privacy. We then asked about how respondents
244 used their VAs as well as any IoT devices or ubicomp technologies they used on a daily basis
245 domestically.

246 *3.1.3 Analyzing the Interviews.* The interviews were transcribed and the transcripts coded using
247 NVivo, a qualitative data analysis package.⁸ The interviews were analyzed using an inductive

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

⁸<http://www.qsrinternational.com/nvivo/nvivo-products>.

Table 1. Interviewee Details

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

AA: Amazon Alexa, GH: Google Home, S: Siri.

process in which the first author conducted multiple passes, discussing the emerging codes after each pass with co-authors. 248 249

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 VAs in their everyday lives. 250 251 252 253 254 255

3.2 Surveys 256

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. 257 258 259 260 261

Table 2. Key Codes for Interview Analysis

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 3. Summary of Survey Results

	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

*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.

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

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

273 The users were asked to answer a question about their geographic location. Since the VA logs
 274 store timestamps in Unix epochs,⁹ these data were used to localize the timestamps from each of
 275 the user logs. We also allowed respondents to provide some free-text responses discussing their
 276 experiences with VAs. The survey took an average of 6 minutes to complete.

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

283 3.3 History Logs: Dataset

284 We used Amazon Mechanical Turk and Reddit to recruit participants to provide us with full device
 285 usage logs from 82 Amazon Alexa users and 88 Google Home owners. We performed the data
 286 collection in a manner similar to the phone book data collection study by Bentley and Chen [7].

⁹<https://www.unixtimestamp.com/index.php>.

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 4th 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 five 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 following:

- 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 term frequency–inverse document frequency (TF–IDF)¹² to the commands to find the terms with the highest scores. 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.

¹⁰<https://alexa.amazon.com>, <https://myactivity.google.com>.

¹¹<https://pandas.pydata.org/>.

¹²Term frequency-inverse document frequency is a score used to find the most important words in a corpus of documents.

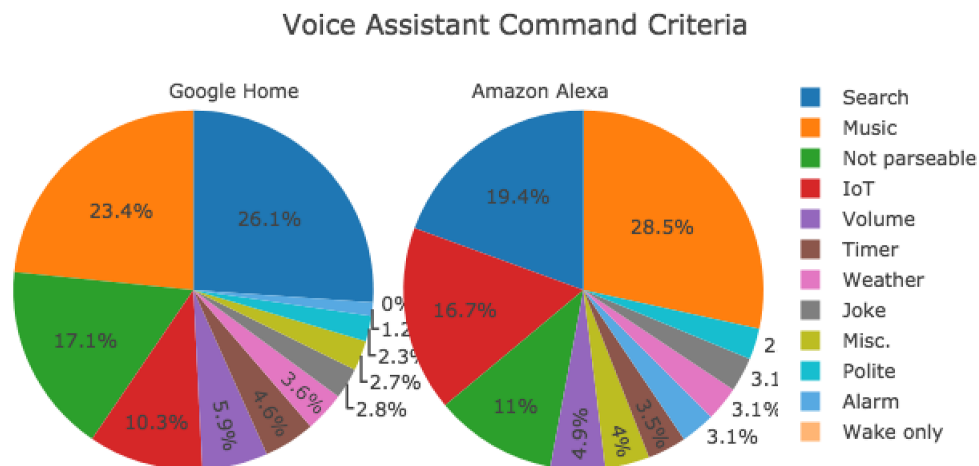


Fig. 1. Breakdown of command categories on Google Home and Amazon Alexa.

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

345 4 RESULTS

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

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

356 Out of the 193,665 Amazon Alexa commands, we found that 51,491 commands consisted only of
 357 wake words like “Amazon,” “Alexa,” “Echo,” and “Computer.” Ford and Palmer [18] have previously
 358 reported that Alexa devices will sometimes spontaneously wake without intentional invocation

Music, Search, and IoT: How People (Really) Use Voice Assistants 17:11

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.

4.1 Music, Media, and Volume 363

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). VA 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. 369
370

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. 374

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. 380
381
382

Figure 2(a) 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

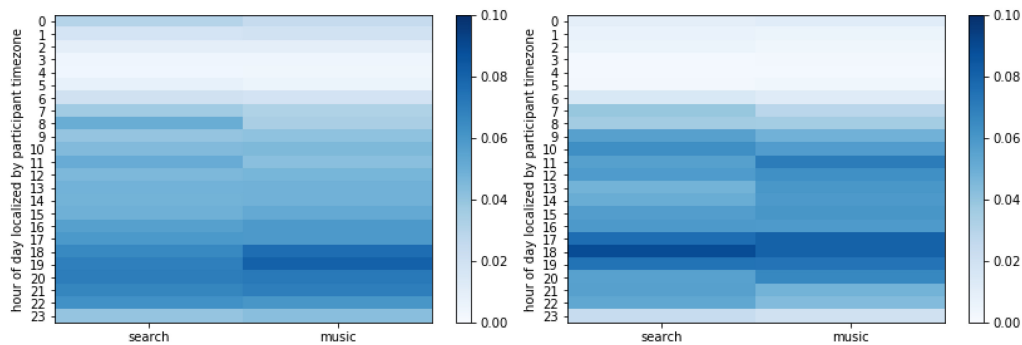
$$Command\ Weight = \frac{\sum_{i=0}^T Music\ Commands}{\sum_{i=0}^T All\ Commands}$$

where T is fixed at 1-hour intervals 386

For Amazon Alexa, the music command was used most heavily between 6 and 10 p.m., while peaking between 6 and 8 p.m. Figure 2(b) 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 p.m. and 9 p.m. 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 VA use to music. Some interviewees indicated that they used their VA to access other media. For example, Jose noted how he used Google Home, along with



(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.

399 Google Chromecast¹³ to operate his Netflix account. Brad discussed how he used Amazon, along
400 with Harmony Hub¹⁴ to control his entertainment center:

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

406 Other interviewees were introduced to smart home devices as they integrated smart speakers
407 (e.g., Sonos) to be used with their VAs. Chuck notes that "the fact that [Alexa] could be auxiliary
408 plugged into a Sonos Play 5 was appealing to me ... They are always expanding their skillset [sic]
409 and there's a big open source community around building integrations to Alexa." Using VAs led
410 users to use smart home devices (e.g., Harmony Hub) for the purpose of interaction with their
411 media environment. Brad describes the layout of VAs in the house thusly

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

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

419 4.2 Search

420 Search or informational queries was the most prevalent use of Google Home (at 26%) and second
421 most prevalent use for Amazon Alexa (at 19.4%). The frequency of search command use was highest

¹³https://store.google.com/product/chromecast_2015.

¹⁴<https://www.logitech.com/en-us/product/harmony-hub>.

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

17:13

for both Amazon Alexa and Google Home was between 5 and 7 p.m. followed by the time between 8 a.m. and noon. 422
423

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, and so on. One respondent notes, 424
425
426
427

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. 428
429

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 2(a) for Amazon Alexa and Figure 2(b) for Google Home. 430
431
432
433
434
435
436

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 437
438

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. 439
440
441

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. 442
443
444
445

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. 446
447
448

Brad compared the search feature for Amazon Alexa and Google Home 449

“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.” 450
451

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. 452
453
454

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 p.m. 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, 455
456
457
458
459

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, whereas my phone? Oh wait, where’s the little dial timer? 460
461
462

463 Timers could also be used to set reminders for users. Table 5 shows an example where a user
464 sets a timer to “remind me to make a smoothie at 11 a.m.” Monique, who has ADHD, said that she
Q3 465 used timers to stay focused on the task at hand.

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

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

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

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

482 Most interviewees also noted that they used VAs as Alarms. The terms used in VA logs as pre-
483 sented in Tables 4 and 5 show the Alarm category includes words like “set” as in set the Alarm
484 and “snooze” when snoozing the alarm when triggered.

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

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

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

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

500 4.4 How Voice Assistants Motivated Home Automation

501 IoT commands were the third most uttered commands for both VAs. IoT commands constitute
502 around 10% of Google Home commands and 16.7% of the Amazon Alexa commands. Both Amazon
503 Alexa and Google Home provide some form of home automation integration. One respondent
504 commented that his VA, an Alexa, provided “many integration points it has with home automation
505 products and account linking abilities with other services makes it a very useful product for me.”
506 The effect of buying a VA was to motivate owners to use these integration points.

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

17:15

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 20%.” Finally, a smaller minority of the commands involved changing the temperature in different parts of the house. For example, “Set kitchen temperature to 76°.” 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

551 devices to a VA. Duke wished he could “see all connected devices without needing to go through a
 552 smarthub.” Harriet also noted that she had some problems integrating Alexa with her IoT devices.
 553 She explains that “Alexa gets confused because there are two different accounts for the same device.
 554 For example, [there might be] two kitchen lamps.” The reason for the duplication usually has to do
 555 with using multiple apps and/or smart hubs to control and integrate IoT devices at home. Harriet
 556 had been using both SmartThings (a Samsung IoT device hub) and Wink (another IoT device hub).
 557 On another note, Brad noted that “it would be great” if the VA could understand the context of
 558 the command. For example, Brad was interested in having the VA better interpret his comments
 559 with relation to his location in the house at the time of issuing the command to Alexa, “when I’m
 560 in the living room and I ask Alexa to shut down the lamp, I want her to shut down the lights in the
 561 living room.” Similarly, another respondent commented that “‘Alexa turn on bedside lamp’ could
 562 mean a different lamp based on who says it.”

563 **4.4.2 Advanced IoT Functionality: Macros and Routines.** Five of the respondents created IoT trig-
 564 gers that can be initiated using Alexa. Throughout the Amazon Alexa command logs, there were
 565 only 338 triggers used throughout the Amazon Alexa logs. A trigger is the command used to initi-
 566 ate an IoT Hub macros. For example, a user might program a “Play Xbox” trigger that would turn
 567 on the TV, Xbox, and stereo, and dim the living room lights. To create or change macros, the user
 568 would have to make updates in the software used to manage their IoT hub.

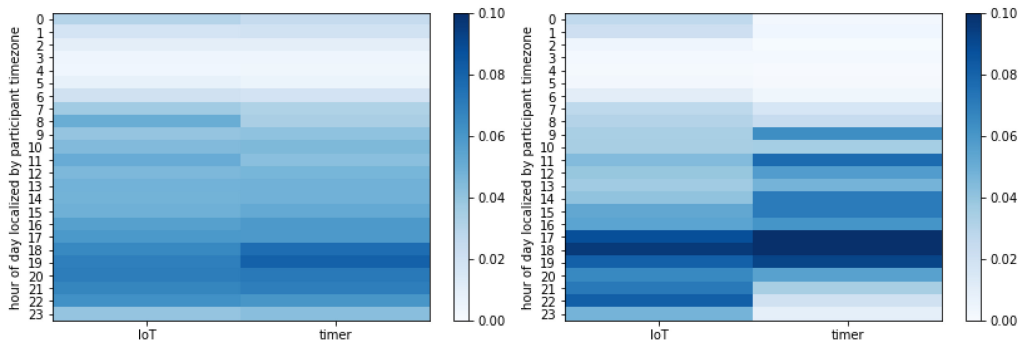
569 “I love Alexa, although there’s lots of things I wish it could do that it can’t (like single commands
 570 to play music and trigger home automation functions)” commented one of the survey respondents
 571 in the free-text question. Similarly, Brad noted that he would like to be able to give Alexa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

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

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

589 **4.5 VAs and Privacy**

590 Amazon and Google have sought to provide history logs for their users in order to enhance their
 591 experience using the VA. Activity logs also provide some transparency and control around data
 592 collection (Alexa History and the Google Activity dashboard). Users are able to view the tran-
 593 scription of audio clips, listen to the audio, see Alexa or Google’s response, and delete items. We
 594 found that survey respondents reported they were aware (69.5%) of the log history, although only
 595 a small percentage reported that they had ever deleted any of their log entries (10.9%). Some of the



(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.

interviewees noted that they used their logs to review their interactions and make sure that 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).

619 4.5.1 *Is She Always Listening?* A survey respondent noted that Alexa sometimes “randomly
620 lights up or is “listening” when I haven’t spoken to her.” Their concern is specific to the control/or
621 lack thereof over when their VA is on/off. Molly, who had an Alexa in the living room, similarly
622 noted that there are topics she would prefer not to discuss around Alexa, like family finances and
623 other issues of a personal nature. Another respondent, Mona explains “I prefer to mute her all
624 the time unless we’re actually using her for something.” Mona expanded on that point by noting
625 that “basically, if we’re having sex we mute Alexa. Just in case [because] sometimes she’ll start
626 blinking...” without a wake word. Mona was referring to a device the couple had in their bedroom.
627 She followed this comment by saying that “[my boyfriend] thinks I’m paranoid” for muting the
628 VA when not in use. This disparity between different home members when contemplating privacy
629 settings have been echoed by other respondents. For example, Brad’s wife was worried that having
630 so many microphones across the house would inevitably result in some privacy invasion. Harriet
631 and Brad both said that that are heavy users of Alexa. They both have more than one VA in different
632 parts of the house. However, Brad and Harriet were both criticized by family members for having
633 too many VAs around their homes. Harriet said that her

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

637 4.5.2 *VA Logs.* Other users had more specific privacy concerns. One such concern centered on
638 the availability of records for their interactions with VAs and the location where these records
639 were stored. For example, one survey respondent said he was, “honestly creeped out that [Alexa]
640 stores so much information I was completely unaware of on a website that’s easy to hack.” This
641 comment was a reaction to the fact that the respondent did not know of the existence of the Alexa
642 History log before he was introduced to it in our study.

643 4.5.3 *Access to Data by Third Party Apps.* John noted that he was concerned about how VAs
644 “reach out to... third party services” when for example asking about the weather. He is critical of
645 the fact that he knows very little about what information is sent to third party services and how
646 these data are stored and protected. He followed this comment by saying that he would rather
647 have “locally hosted” systems where he can be in better control of his data. Similarly, one of our
648 survey respondents suggested that he wished for a “an open-source locally hosted alternative” VA
649 for domestic use.

650 **5 DISCUSSION**

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

656 **5.1 RQ1: What are the Daily Uses of VAs?**

657 With our analysis of Alexa and Google Home History and Activity data, we have a more con-
658 crete and accurate understanding of how people are using their VAs (especially compared to self-
659 reported usage). We found that the three main uses for both Google Home and Amazon Alexa are
660 (1) music, (2) hands free search, and (3) IoT control, primarily turning lights on and off. We also
661 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 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 while 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 while 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,

707 bedroom, living room, and so on.) mostly to turn lights on and off. While some users found such
708 use convenient, this idea was a vague one related to making the home a “modern” or “smart home.”
709 This echoes the findings of Mennicken and Huang as they studied the introduction of IoT devices
710 in domestic spaces outside of laboratory settings [35].

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

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

726 *5.2.1 Spatial and Temporal Contextualization.* Spatial contextualization refers to the capacity
727 of the VA to recognize where the user is physically at any particular point in time. If a user wants
728 to control an IoT device in the living room while in the living room, then the VA should understand
729 that the user is attempting to control the IoT device in the living room, unless otherwise specified
730 by the user. Similarly, Guha et al. in their design recommendations argued for the importance of
731 geographical contextualization of the data used by the VA [19]. In their case, this contextualization
732 required the use of the GPS coordinates of the user.

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

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

747 *5.2.2 Dynamic Instructions.* When using VAs to control domestic IoT devices, users indicated
748 that they wanted to dynamically control IoT actions via what two of the users termed “macros.”
749 Macros would allow users to control a number of different IoTs in relation to a specific activity.
750 For example, if the user is leaving the house, she might want to turn off lights in the house, close
751 the garage door, and reduce the temperature on the thermostat.

752 At this point, the only way for users to create macros is by programming them through IoT hubs.
753 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 p.m. until the morning alarm.

Our respondents made it clear that they did not know what information was shared with 3rd 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

800 fraction of those who knew of the existence of the logs edited them for privacy concerns. This
801 finding echos findings by Lua et al. [26] indicating that while users might know of the logs, they
802 might find accessing and editing them too cumbersome.

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

815 Some of the respondents who did have privacy concerns were most worried about the fact
816 that their speech is being processed remotely. VA producers can ameliorate the users' con-
817 cerns by providing detailed information about when and with whom these data will be shared
818 [60].

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

823 6 LIMITATIONS AND FUTURE WORK

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

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

838 While we articulated the broad command categories of VA use, future work could focus on the
839 effects of current VA uses on future use patterns. Another area to investigate is the adoption of VAs
840 by different user profiles. For example, can the current use of VAs predict future use of VAs by the
841 user? Does the use of Harmony hubs, which have some IoT characteristics, result in the increased
842 use of IoT devices? As the use of VAs like Siri and Google assistant (usually on cell phones) increase
843 [43], how is the use of these technologies affecting the way users think of VAs at home? How does
844 it affect the way they decide whether to adopt VAs at home or not? How does it affect their privacy
845 concerns when using VAs?

7 CONCLUSIONS

846

As 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 VAs. 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.

847

848

849

850

851

852

853

854

855

APPENDIX A

856

Table 4. This Table Shows Amazon Alexa Command Categories Along with Highest Frequency Words and High-Score TF-IDF Scores for Each Category

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

We also provide a few examples for each category.

857

Table 5. This Table Shows Google Home Command Categories Along with Highest Frequency Words and High-Score TF-IDF Scores 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

858 We also provide a few examples for each category.

859 A.1 An Example of Command Category Iteration

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

```

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

872 After the first few iterations, we found that there are other commands that, while using some of
 873 the terms in the regular expression above, do not relate to playing music. For example, we found

that some of the users were “playing” 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 881
not_music_criteria= 882
r'^ (?!Text not available. Click to play recording.|news|jeopardy|stop the alarm.)*$' 883
```

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 891
#the new miscellaneous data frame, df_miscellaneous will contain all commands not 892
#picked in any of the nine command categories identified earlier 893
df_miscellaneous = df {~df.command.isin(all_commands)} 894
```

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, 1895–1904. DOI : <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, 229–238. DOI : <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. DOI : <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. DOI : <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. DOI : <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, 2749–2758. DOI : <https://doi.org/10.1145/2702123.2702182>
- [8] Frank R. Bentley, Nediya 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, 1092–1099. DOI : <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, 1390–1401. DOI : <https://doi.org/10.1145/2818048.2819928>

- 924 [11] Rodney A. Brooks. 1997. The intelligent room project. In *Proceedings of the 2nd International Conference on Cognitive*
925 *Technology (CT'97)*. IEEE, 271–278.
- 926 [12] A. J. Bernheim Brush, Bongshin Lee, Ratul Mahajan, Sharad Agarwal, Stefan Saroiu, and Colin Dixon. 2011. Home
927 automation in the wild: Challenges and opportunities. In *Proceedings of the SIGCHI Conference on Human Factors in*
928 *Computing Systems (CHI'11)*. ACM, New York, NY, 2115–2124. DOI: <https://doi.org/10.1145/1978942.1979249>
- 929 [13] Yi Cheng, Kate Yen, Yeqi Chen, Sijin Chen, and Alexis Hiniker. 2018. Why doesn't it work?: Voice-driven interfaces
930 and young children's communication repair strategies. In *Proceedings of the 17th ACM Conference on Interaction Design*
931 *and Children (IDC'18)*. ACM, New York, NY, 337–348. DOI: <https://doi.org/10.1145/3202185.3202749>
- 932 [14] Andy Crabtree and Tom Rodden. 2004. Domestic routines and design for the home. *Comput. Supported Coop. Work*
933 *13*, 2 (April 2004), 191–220. DOI: <https://doi.org/10.1023/B:COU.0000045712.26840.a4>
- 934 [15] Andy Crabtree and Peter Tolmie. 2016. A day in the life of things in the home. In *Proceedings of the 19th ACM*
935 *Conference on Computer-Supported Cooperative Work & Social Computing (CSCW'16)*. ACM, New York, NY, 1738–
936 1750. DOI: <https://doi.org/10.1145/2818048.2819954>
- 937 [16] Wenrui Diao, Xiangyu Liu, Zhe Zhou, and Kehuan Zhang. 2014. Your voice assistant is mine: How to abuse speakers
938 to steal information and control your phone. In *Proceedings of the 4th ACM Workshop on Security and Privacy in*
939 *Smartphones & Mobile Devices (SPSM'14)*. ACM, New York, NY, 63–74. DOI: <https://doi.org/10.1145/2666620.2666623>
- 940 [17] Colin Dixon, Ratul Mahajan, Sharad Agarwal, A. J. Brush, Bongshin Lee, Stefan Saroiu, and Victor Bahl. 2010. The
941 home needs an operating system (and an app store). In *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics*
942 *in Networks (Hotnets-IX)*. ACM, New York, NY, 18:1–18:6. DOI: <https://doi.org/10.1145/1868447.1868465>
- 943 [18] Marcia Ford and William Palmer. 2018. Alexa, are you listening to me? An analysis of Alexa voice service network
Q4 944 traffic. *Pers. Ubiquitous Comput.* (Jun. 2018), 1–13. DOI: <https://doi.org/10.1007/s00779-018-1174-x>
- 945 [19] Ramanathan Guha, Vineet Gupta, Vivek Raghunathan, and Ramakrishnan Srikant. 2015. User modeling for a personal
946 assistant. In *Proceedings of the 8th ACM International Conference on Web Search and Data Mining (WSDM'15)*. ACM,
947 New York, NY, 275–284. DOI: <https://doi.org/10.1145/2684822.2685309>
- 948 [20] Jason I. Hong and James A. Landay. 2004. An architecture for privacy-sensitive ubiquitous computing. In *Proceedings*
949 *of the 2nd International Conference on Mobile Systems, Applications, and Services (MobiSys'04)*. ACM, New York, NY,
950 177–189. DOI: <https://doi.org/10.1145/990064.990087>
- 951 [21] S. S. Intille. 2002. Designing a home of the future. *IEEE Perv. Comput.* 1, 2 (April 2002), 76–82. DOI: <https://doi.org/10.1109/MPRV.2002.1012340>
- 952 [22] Joseph 'Jofish' Kaye, Joel Fischer, Jason Hong, Frank R. Bentley, Cosmin Munteanu, Alexis Hiniker, Janice Y. Tsai, and
953 Tawfiq Ammari. 2018. Panel: voice assistants, UX design and research. In *Proceeding of the Extended Abstracts of the*
954 *2018 CHI Conference on Human Factors in Computing Systems (CHI EA'18)*. ACM, New York, NY, panel01:1–panel01:5.
955 DOI: <https://doi.org/10.1145/3170427.3186323>
- 956 [23] Julie A. Kientz, Shwetak N. Patel, Brian Jones, Ed Price, Elizabeth D. Mynatt, and Gregory D. Abowd. 2008. The
957 Georgia tech aware home. In *Proceeding of the Extended Abstracts on Human Factors in Computing Systems (CHI*
958 *EA'08)*. ACM, New York, NY, 3675–3680. DOI: <https://doi.org/10.1145/1358628.1358911>
- 959 [24] Julia Kiseleva, Kyle Williams, Jiepu Jiang, Ahmed Hassan Awadallah, Aidan C. Crook, Imed Zitouni, and Tasos
960 Anastasakos. 2016. Understanding user satisfaction with intelligent assistants. In *Proceedings of the 2016 ACM*
961 *on Conference on Human Information Interaction and Retrieval (CHIIR'16)*. ACM, New York, NY, 121–130. DOI:
962 <https://doi.org/10.1145/2854946.2854961>
- 963 [25] Predrag Klasnja, Sunny Consolvo, Tanzeem Choudhury, Richard Beckwith, and Jeffrey Hightower. 2009. Explor-
964 ing privacy concerns about personal sensing. In *Pervasive Computing (Lecture Notes in Computer Science)*, Hideyuki
965 Tokuda, Michael Beigl, Adrian Friday, A. J. Bernheim Brush, and Yoshito Tobe (Eds.). Springer, Berlin Heidelberg,
966 176–183.
- 967 [26] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, are you listening?: Privacy perceptions, con-
968 cerns and privacy-seeking behaviors with smart speakers. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (2018), 102.
- 969 [27] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhammedi, S. A. Zhang, Norman Sadeh, Alessandro
970 Acquisti, and Yuvraj Agarwal. 2016. Follow my recommendations: A personalized privacy assistant for mobile app
971 permissions. In *Proceeding of the Symposium on Usable Privacy and Security*. [https://www.usenix.org/system/files/](https://www.usenix.org/system/files/conference/soups2016/soups2016-paper-liu.pdf)
972 [conference/soups2016/soups2016-paper-liu.pdf](https://www.usenix.org/system/files/conference/soups2016/soups2016-paper-liu.pdf).
- 973 [28] Silvia Lovato and Anne Marie Piper. 2015. "Siri, is this you?": Understanding young children's interactions with voice
974 input systems. In *Proceedings of the 14th International Conference on Interaction Design and Children (IDC'15)*. ACM,
975 New York, NY, 335–338. DOI: <https://doi.org/10.1145/2771839.2771910>
- 976 [29] Ewa Luger and Tom Rodden. 2013. An informed view on consent for UbiComp. In *Proceedings of the ACM Interna-*
977 *tional Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'13)*. ACM, New York, NY, 529–538. DOI:
978 <https://doi.org/10.1145/2493432.2493446>
- 979

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

17:27

- [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 Conference on Human Factors in Computing Systems (CHI’16)*. ACM, New York, NY, 5286–5297. DOI : <https://doi.org/10.1145/2858036.2858288> 980–982
- [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, 2273–2286. DOI : <https://doi.org/10.1145/2998181.2998218> 983–985
- [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. 986–988
- [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 Conference on Human Factors in Computing Systems (CHI’17)*. ACM, New York, NY, 5197–5207. DOI : <https://doi.org/10.1145/3025453.3025735> 989–991
- [34] Michael F. McTear. 2002. Spoken dialogue technology: Enabling the conversational user interface. *ACM Comput. Surv.* 34, 1 (Mar. 2002), 90–169. DOI : <https://doi.org/10.1145/505282.505285> 992–993
- [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. DOI : https://doi.org/10.1007/978-3-642-31205-2_10 994–996
- [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 International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp’16)*. ACM, New York, NY, 670–673. DOI : <https://doi.org/10.1145/2968219.2968345> 997–1000
- [37] John Moore, Gerd Kortuem, Andrew Smith, Niaz Chowdhury, Jose Cavero, and Daniel Gooch. 2016. DevOps for the urban IoT. In *Proceedings of the 2nd International Conference on IoT in Urban Space (Urb-IoT’16)*. ACM, New York, NY, 78–81. DOI : <https://doi.org/10.1145/2962735.2962747> 1001–1003
- [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. DOI : https://doi.org/10.1007/978-3-319-07863-2_32 1004–1007
- [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, 471–480. DOI : <https://doi.org/10.1145/2639189.2639239> 1008–1011
- Q5 [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. 1012–1013
- [41] Clifford Nass, Youngme Moon, B. J. Fogg, Byron Reeves, and D. Christopher Dryer. 1995. Can computer personalities be human personalities? *Int. J. Hum.-Comput. Studies* 43, 2 (Aug. 1995), 223–239. DOI : <https://doi.org/10.1006/ijhc.1995.1042> 1014–1016
- [42] Helen Nissenbaum. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press. Google-Books-ID: _NN1uGn1Jd8C. 1017–1018
- [43] Kenneth Olmstead. 2017. Nearly half of Americans use digital voice assistants, mostly on their smartphones. <http://www.pewresearch.org/fact-tank/2017/12/12/nearly-half-of-americans-use-digital-voice-assistants-mostly-on-their-smartphones/>. 1019–1021
- [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>. 1022–1024
- [45] Leysia Palen and Paul Dourish. 2003. Unpacking privacy for a networked world. ACM, 129–136. DOI : <https://doi.org/10.1145/642611.642635> 1025–1026
- [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, 455–464. DOI : <https://doi.org/10.1145/1394445.1394494> 1027–1029
- [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. DOI : <https://doi.org/10.1145/2702123.2702325> 1030–1032
- [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 Conference on Computer Supported Cooperative Work and Social Computing (CSCW’17)*. ACM, New York, NY, 207–219. DOI : <https://doi.org/10.1145/2998181.2998298> 1033–1034

- 1036 [49] Alisha Pradhan, Kanika Mehta, and Leah Findlater. 2018. Accessibility came by accident: Use of voice-controlled intel-
1037 ligent personal assistants by people with disabilities. In *Proceedings of the Conference on Human Factors in Computing*
1038 *Systems*. ACM, 459.
- 1039 [50] Amanda Purington, Jessie G. Taft, Shruti Sannon, Natalya N. Bazarova, and Samuel Hardman Taylor. 2017. "Alexa is
1040 my new BFF": Social roles, user satisfaction, and personification of the Amazon echo. In *Proceedings of the Conference*
1041 *Extended Abstracts on Human Factors in Computing Systems (CHI EA'17)*. ACM, New York, NY, 2853–2859. DOI :
1042 <https://doi.org/10.1145/3027063.3053246>
- 1043 [51] Juan Ramos. 2003. Using tf-idf to determine word relevance in document queries. In *Proceedings of the 1st Instructional*
1044 *Conference on Machine Learning*, Vol. 242. 133–142.
- 1045 [52] Stuart Reeves and Barry Brown. 2016. Embeddedness and sequentiality in social media. In *Proceedings of the 19th ACM*
1046 *Conference on Computer-Supported Cooperative Work & Social Computing (CSCW'16)*. ACM, New York, NY, 1052–1064.
1047 DOI : <https://doi.org/10.1145/2818048.2820008>
- 1048 [53] Jennifer A. Rode, Eleanor F. Toye, and Alan F. Blackwell. 2005. The domestic economy: A broader unit of analysis for
1049 end user programming. In *Proceedings of the Extended Abstracts on Human Factors in Computing Systems (CHI EA'05)*.
1050 ACM, New York, NY, 1757–1760. DOI : <https://doi.org/10.1145/1056808.1057015>
- 1051 [54] Benjamin Romano. 2017. Managing the internet of things. In *Proceedings of the 2017 ACM SIGCSE Technical Sympo-*
1052 *sium on Computer Science Education (SIGCSE'17)*. ACM, New York, NY, 777–778. DOI : [https://doi.org/10.1145/3017680.](https://doi.org/10.1145/3017680.3022452)
1053 [3022452](https://doi.org/10.1145/3017680.3022452)
- 1054 [55] Xin Rong, Adam Fourney, Robin N. Brewer, Meredith Ringel Morris, and Paul N. Bennett. 2017. Managing uncertainty
1055 in time expressions for virtual assistants. In *Proceedings of the Conference on Human Factors in Computing Systems*
1056 *(CHI'17)*. ACM, New York, NY, 568–579. DOI : <https://doi.org/10.1145/3025453.3025674>
- 1057 [56] Gerard Salton and Christopher Buckley. 1988. Term-weighting approaches in automatic text retrieval. *Inf. Process.*
1058 *Manag.* 24, 5 (1988), 513–523.
- 1059 [57] Nicole Shechtman and Leonard M. Horowitz. 2003. Media inequality in conversation: How people behave differently
1060 when interacting with computers and people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing*
1061 *Systems (CHI'03)*. ACM, New York, NY, 281–288. DOI : <https://doi.org/10.1145/642611.642661>
- 1062 [58] Peter Tolmie, Andy Crabtree, Stefan Egglestone, Jan Humble, Chris Greenhalgh, and Tom Rodden. 2010. Digital
1063 plumbing: The mundane work of deploying UbiComp in the home. *Personal and Ubiquitous Computing* 14, 3 (Apr.
1064 2010), 181–196. DOI : <https://doi.org/10.1007/s00779-009-0260-5>
- 1065 [59] Peter Tolmie, Andy Crabtree, Tom Rodden, Chris Greenhalgh, and Steve Benford. 2007. Making the home net-
1066 work at home: Digital housekeeping. In *ECSCW 2007*. Springer, London, 331–350. DOI : [https://doi.org/10.1007/](https://doi.org/10.1007/978-1-84800-031-5_18)
1067 [978-1-84800-031-5_18](https://doi.org/10.1007/978-1-84800-031-5_18)
- 1068 [60] Janice Y. Tsai, Patrick Kelley, Paul Drielsma, Lorrie Faith Cranor, Jason Hong, and Norman Sadeh. 2009. Who's viewed
1069 you?: The impact of feedback in a mobile location-sharing application. In *Proceedings of the SIGCHI Conference on*
1070 *Human Factors in Computing Systems (CHI'09)*. ACM, New York, NY, 2003–2012. DOI : [https://doi.org/10.1145/1518701.](https://doi.org/10.1145/1518701.1519005)
1071 [1519005](https://doi.org/10.1145/1518701.1519005)
- 1072 [61] Giorgio Vassallo, Giovanni Pilato, Agnese Augello, and Salvatore Gaglio. 2010. Phase coherence in conceptual
1073 spaces for conversational agents. In *Semantic Computing*. Wiley-Blackwell, 357–371. DOI : [https://doi.org/10.1002/](https://doi.org/10.1002/9780470588222.ch18)
1074 [9780470588222.ch18](https://doi.org/10.1002/9780470588222.ch18)
- 1075 [62] Sergey Volokhin and Eugene Agichtein. 2018. Understanding music listening intents during daily activities with
1076 implications for contextual music recommendation. In *Proceedings of the on Human Information Interaction & Retrieval*
1077 *(CHIIR'18)*. ACM, New York, NY, 313–316. DOI : <https://doi.org/10.1145/3176349.3176885>
- 1078 [63] Rick Wash, Emilee Rader, and Chris Fennell. 2017. Can people self-report security accurately?: Agreement between
1079 self-report and behavioral measures. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI'17)*.
1080 ACM, New York, NY, 2228–2232. DOI : <https://doi.org/10.1145/3025453.3025911>
- 1081 [64] Thomas Zachariah, Noah Klugman, Bradford Campbell, Joshua Adkins, Neal Jackson, and Prabal Dutta. 2015. The
1082 internet of things has a gateway problem. In *Proceedings of the 16th International Workshop on Mobile Computing*
1083 *Systems and Applications (HotMobile'15)*. ACM, New York, NY, 27–32. DOI : <https://doi.org/10.1145/2699343.2699344>

1084 Received July 2018; revised January 2019; accepted February 2019

Author Queries

- [Q1](#): AU: Please provide complete mailing and email address.
- [Q2](#): AU: Please check all the sections headings are okay as set.
- [Q3](#): AU: Please define the term “ADHD” at its first occurrence in the text.
- [Q4](#): AU: Please provide volume number in Ref. [18].
- [Q5](#): AU: Please provide complete details in Ref. [40]