

“All Rise for the AI Director”: Eliciting Possible Futures of Voice Technology through Story Completion

Julia Cambre, Samantha Reig, Queenie Kravitz, Chinmay Kulkarni

Human-Computer Interaction Institute, Carnegie Mellon University

{jcambre, sreig, qkravitz, chinmayk}@cs.cmu.edu

ABSTRACT

How might the capabilities of voice assistants several decades in the future shape human society? To anticipate the space of possible futures for voice assistants, we asked 149 participants to each complete a story based on a brief story stem set in the year 2050 in one of five different contexts: the home, doctor’s office, school, workplace, and public transit. Story completion as a method elicits participants’ visions of possible futures, unconstrained by their understanding of current technological capabilities, but still reflective of current sociocultural values. Through a thematic analysis, we find these stories reveal the extremes of the capabilities and concerns of today’s voice assistants—and artificial intelligence—such as improving efficiency and offering instantaneous support, but also replacing human jobs, eroding human agency, and causing harm through malfunction. We conclude by discussing how these speculative visions might inform and inspire the design of voice assistants and other artificial intelligence.

Author Keywords

voice assistant; conversational user interface; CUI; story completion; speculative design

CCS Concepts

•Human-centered computing → Human computer interaction (HCI);

INTRODUCTION

A lot has changed since the first inclusion of the term ‘voice assistant’ into polite conversation back in 2010. The chirpy voice-detained assistants who were once constrained to small pocket devices left perching on lounge fireplaces or bedside tables are no longer quite the small fry they once were. How would you manage without your walking, talking, miming, moving HomeInstalled SatNav 3000? Well, you wouldn’t, Sam thought. (S119–H)

This excerpt highlights one of the 149 different visions of futuristic voice technology that participants contributed in our study, offering a rich and varied range of perspectives on what individuals today believe voice assistants might become

Permission to make digital or hard copies of part or all 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 third-party components of this work must be honored. For all other uses, contact the owner/author(s).

DIS '20, July 06–10, 2020, Eindhoven, NL

© 2020 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6974-9/20/07.

DOI: <http://dx.doi.org/10.1145/3357236.3395479>

several decades from now. As the narrative unfolds, it depicts a future world where assistants cater to a human’s every desire as the ultimate butler might, but are also accorded “equal mental status in the eyes of the law”. Other stories suggest instead a future where assistants are the orchestrators of human action, at times taking away both their jobs and their agency. What can we learn from these diverse visions of the future?

Fiction is particularly useful in informing design as a method of speculating on *possible futures*—scenarios that allow their audience to “wonder about how things could be” [23, pg. 3]. As Dourish and Bell write, “science fiction does not merely anticipate but actively shapes technological futures through its effect on the collective imagination. At the same time, science fiction in popular culture provides a context in which new technological developments are understood” [21, pg. 769]. This idea of using fiction—both science fiction and other forms of fiction—to inform design has been furthered and popularized by a number of other academics and practitioners, who argue that fictions can inspire and inform prototypes and even research outcomes before they come to fruition [7, 53]. Fictions can also help to articulate the potential implications for a technology’s adoption, anticipating the downstream effects that it may have on society [28]. However, because “social and cultural themes of the times also surface as science fiction conceptualizes a future world” [45, pg. 39], and fiction is not tied to actual capabilities of technology, it is best seen as an informative probe into potential futures, and not a prediction of what the future *will* be.

Because voice assistants have grown enormously popular, but still have a complicated role at present, taking a futures-oriented approach can be particularly meaningful for their design. As of late 2019, estimates suggest that nearly a quarter of adults in the United States (60 million) own a smart speaker [38], and Google reports that over half a billion people use their Google Assistant at least once a month [32]. Apple’s Siri, Amazon’s Alexa, Microsoft’s Cortana, Baidu’s DuerOS and many other voice assistants have also gained widespread adoption [26, 39]. In recent years, the physical embodiment of voice assistants has diversified as well, with assistants embedded not just in smartphones, but in smart speakers, cars, and a range of other smart home devices [10]. At the same time, voice assistants pose certain challenges. From a usability perspective, users often have inflated expectations of capabilities [19, 30, 40] and struggle to know what to say [18, 37]. Recent studies also point to privacy concerns and perceptions

that devices are always listening [19, 27], as well as social concerns around how the default gendering of many popular voice assistants may perpetuate gender stereotypes [10, 52, 54].

Taken together, these studies suggest that voice assistants occupy a complicated position: part-speech-activated interface, part-artificial intelligence, and part-complex social actors with difficult-to-grasp or ambiguous capabilities that can spark concerns about stereotyping, privacy, and more. Given the complicated picture that has emerged from their current capabilities, it becomes crucially important to reflect on what voice assistants *might become*.

With this in mind, this paper asks: what possible futures might emerge around voice assistants several decades in the future? To anticipate these possible futures, it uses *story completion* as a design method. In this study, we provided participants with the opening to a story set 30 years in the future in one of five different contexts, and asked them to complete the story however they imagined. While these stories were situated in the distant future of the year 2050, echoing Dourish and Bell [21] in their reading of science fiction, we take these narratives as also reflecting back in informative ways on the present: “An account of ‘how we shall live’ is inherently grounded in assumptions about the problems and opportunities of the time at which it is written” [21, pg. 770].

Through thematic analysis and iterative coding of the 149 stories, we identified several themes within these voice assistant futures. Following Dourish and Bell, we see the corpus of stories as reflecting the extremes of the capabilities and concerns of today’s voice assistants, as well as the trade-offs implied by these extremes. The futures that participants envisioned depicted integration of different technologies; blurred boundaries between human and artificial intelligence (e.g., brain-computer interfaces); agent roles ranging from friends to foes, and slaves to dictators; and various visions of efficiency, some of which suggest the erosion of human agency.

We offer our analyses of these stories not as scientific fact drawn from absolute truths of the present, but as a collage that, when viewed as a whole, constitutes a snapshot of possible futures from the vantage point of the present. We conclude with some of the many open questions these stories spark, and what they mean for the design of voice assistants today.

BACKGROUND

This paper draws upon empirical work on current perceptions of voice assistants and the methods of speculative design.

Voice assistants today: functional but flawed

Corresponding to a rise in voice assistant popularity, there has been a surge of interest within the HCI community in studying voice agents and other speech technology [15]. These studies describe how people use and perceive of today’s common commercial assistants such as Siri and Amazon Alexa. For example, analysis of log data suggest that households use assistants most often for a relatively small set of tasks: entertainment (e.g., playing music, telling a joke), hands-free support (e.g., setting a timer), and controlling IoT devices [2, 47]. However, qualitative studies reveal that users often imagine voice-based

assistants to be more capable, technically and socially, than they truly are, leading to mixed experiences.

In particular, users often anthropomorphize voice assistants, in large part due to human-like names, voices, and personalities the assistants are designed to convey [15, 19, 30], and interact with them as though they were social actors. For example, Purtington et al. [44] found that users who referred to the Amazon Alexa using gendered personal pronouns (e.g., she/her) were more likely to engage socially with the assistant compared to those who used object pronouns (e.g., it/its). However, voice assistants today fall far short of users’ expectations of human-like intelligence, leading to disillusionment or non-use over time as a result [14, 19, 22, 40]. This has led some researchers, such as Luger and Sellen, to suggest voice assistants today behave at best as a “*really bad [personal assistant]*” [30].

In addition to mismatched expectations, technical challenges still pervade today’s voice assistants. Voice assistants routinely mis-recognize speech [37], particularly in contexts with difficult to transcribe content like names of musical artists [50] or scientific vocabulary [11]. Voice as an interaction modality also raises unique challenges over graphical user interfaces (GUIs): speech interfaces can impose higher demands on cognitive load [48, 57] and often suffers from poor discoverability [18, 37] as users are not aware of what commands the system can support. Further still, different populations raise different design needs and usability challenges, such as older adults [42], children [5, 29] and settings where multiple people interact simultaneously with an assistant [41].

As a response to these challenges, researchers have recently introduced new guidelines for speech interfaces, suggesting new best practices or heuristics specific to voice and conversational design [35, 36], and calling attention to specific dimensions of speech interfaces, such as how the particular voice given to a system shapes the user experience [3, 10, 34, 51].

Recent efforts have also scrutinized emerging and problematic trends in how voice assistant design might shape societal norms. For example, designers of popular commercial voice assistants have used feminine personas by default in most markets [54], a misguided decision that perpetuates harmful gender stereotypes [10, 52].

This prior work on voice assistants situates the state of voice at present, in contrast with the visions put forward in this paper about what voice assistants might become in the future.

Using fiction as a method of informing design

The HCI and design communities have used fiction in several forms. Perhaps the most iconic is *design fiction*. Following Bleeker [6], “design fiction is a mix of science fact, design and science fiction” that invites designers to reflect on and critique possible future worlds around technology, and creates “stories that speculate about new, different, distinctive social practices that assemble around and through these objects” [6]. Design fiction has been applied to probe a diverse range of topics, some of which tie in to the focus and findings of our work. For example, Wong et al. [55] created a design fiction about brain-computer interfaces (BCIs) to speculate about a potential scenario in which BCIs use unconscious responses

from users to help train a sorting algorithm. Even closer to our topic of study, Søndergaard and Hansen's [52] recent design fictions critique voice assistant technology from a feminist HCI perspective, specifically considering the consequences of how today's voice assistants are gendered in stereotypically feminine ways by default. Design fictions also consider different contextual timescales, with some taking a "historical" view informed by the past [8] or set in the near future [49], whereas others have considered "a time point that is simultaneously both proximal and distant" [4] set twenty-five years in the future—roughly the timescale we apply in this study.

Participatory forms of design fiction

Typically, design researchers author design fictions. While participants can play a role in the process as the fiction's audience, their involvement is more indirect. By contrast, other forms of design fiction are intentionally participatory, directly involving the participant in constructing or co-constructing the fiction [1, 13, 31, 43], leading to unique design insights. For example, Prost et al. [43] conducted workshops in which participants collaboratively wrote a design fiction about sustainability practices in the home. They found that the fictional frame revealed new design dimensions such as user empowerment that researchers had not previously considered, and shifted focus away from immediate practical concerns like saving money, and towards broader societal themes [43].

Other participatory techniques employ a first-person approach in which participants themselves are characters, but in fictionalized circumstances. One such related technique is *fictional inquiry*, which begins with partially fictional contexts (e.g., artifacts and scenarios) constructed by the designer, and asks participants to imagine and enact a desirable future based on the presented plot [20]. The goal of the fictional inquiry is specifically as a reframing tool, and provides more constraints than other methods by stipulating that the futures participants imagine be desirable ones.

Story completion

In this paper, we adopt story completion as a design method. In story completion tasks, participants are provided with a brief, deliberately ambiguous "stem" (or story starter) of a few sentences that establishes the opening of a narrative, and are asked to write the rest of the story [16]. The method has its roots in the field of psychology, and is often used to understand participants' views around sensitive or stigmatized subjects such as infidelity [17, 25] and body hair removal practices [24]. The HCI and Design communities have only recently adopted story completion as a method. Most notably, Wood et al. [56] analyzed 45 stories gathered through story completion to understand "cultural ideals" around virtual reality pornography. By using story completion as a method, the authors note that they "were able to explore the moral dimensions of new and emerging technologies [...] but also explore how a potentially problematic medium could be reimagined" [56]. While voice assistants do not carry the same potential taboos as virtual reality pornography might, we nevertheless see parallels in the relative novelty of the technologies, and in their potential for inspiring both positive and negative outcomes in the future.

METHODS

This study uses *story completion* as a method, following from Wood et al. [56] in applying story completion to HCI questions, and in particular as a means of understanding the design space that non-experts envision for an emerging technology.

Materials: Story prompt development

Crafting well-defined stems requires careful attention to ensure that participants are constrained by the parameters of interest (e.g., the context of the narrative), while leaving other details open to their interpretation and imagination [9, 16]. Within the space of voice assistant interactions, we were interested in how manipulating the *scenario* of the story and *number of main characters* would shape the nature of stories that participants constructed, while also fixing the *temporal setting* of the story in the distant future (the year 2050).

Scenario: To probe at how different contexts of use—all of which entail unique social dynamics, privacy concerns, and more—shape the interactions that users describe, we created five story stems, situated in the contexts of the *home*, *doctor's office*, *school*, *workplace*, and *public transit*.

Number of characters: We also varied the number of characters present within the scene. Prior work suggests that interactions between users and agents differs considerably between single and multi-user contexts [12, 41]. The story stems we constructed therefore imply different numbers of characters (named protagonists, secondary characters, and groups of people) to draw out these differences.

Temporal setting: We chose to situate the stories in a more distant future (the year 2050), approximately 30 years from the time of the study. This temporal distance from the present gives participants license to think creatively about the state of technology and society, and avoids overly constraining their thinking based on how voice assistants currently function.

With these parameters in mind, we created the following five final story stems:

Home	It's the year 2050, several decades from now. Sam gets home from work and enters the living room, where Jamie is sitting. The voice assistant...
Doctor	It's the year 2050, several decades from now. Sam is at the doctor's office with Sam's partner, Jamie, for a yearly checkup. After Sam signs in, the voice assistant...
School	It's the year 2050, several decades from now. Sam is about to begin 9th grade at a new high school, and is anxious to meet new classmates. When Sam walks into the classroom, the voice assistant...
Work	It's the year 2050, several decades from now. Sam walks into a conference room at work, where the rest of the team has already gathered for their weekly meeting. The voice assistant...
Transit	It's the year 2050, several decades from now. Sam and Jamie are taking the subway back home from work. The voice assistant...

Several other features of the story stems are worth noting. First, to avoid priming effects, we attempted to keep story stems as similar as possible in length. All stories begin with the same sentence to establish temporal setting, and end with the same phrase to focus the participant's attention on the

voice assistant's role in the story. In all stems, we use the names Sam and Jamie for the main characters, as both names are gender ambiguous. Several aspects of the story frames differ as well, such as the nature of the relationship between Sam and Jamie (stated as partners at the doctor, but ambiguous elsewhere) and age of the characters (implied as being in the ninth grade in the school context, but ambiguous elsewhere). Taken together, we see both the similarities and differences between these story stems as helping to elicit a diverse—yet still meaningfully comparable—set of potential narratives.

Procedure: Eliciting stories by survey

To elicit stories, we created a Qualtrics-based survey consisting of three parts. After consenting to participate in the study, participants were randomly assigned to one of the five story scenario conditions (between-subjects; one story per participant). In the first part of the study, participants were introduced to the storywriting task, and asked to spend 10 minutes completing the story established by their assigned prompt in a large text field. The instructions for this portion of the task were adapted with minor edits from prior work on story completion [9, pg. 57] and encouraged participants to use their creativity, reassuring them that there were no “right or wrong” answers. To focus attention on the voice assistant, we asked that they give the assistant a name, and that they elaborate on the context (or “scene”) of this interaction. Because stories were authored through an online survey, participants could spend longer than 10 minutes writing if they wished.

Upon submitting their story, we reproduced the story text on the next screen and asked participants to answer the question, “Consider the story you just wrote: Would you like to live in this future?” on a five-point Likert item question from “Definitely” to “Definitely Not”. Once they selected an answer for this question, we prompted them to “Please briefly explain your choice” through an open text field.

In the third part of the survey, we asked participants to provide relevant demographic information including age and gender. We also included questions about voice assistant familiarity and use, as we anticipated that experience with voice assistants might affect the stories that participants generated.

Participants were recruited through the Prolific crowd work platform in mid-October 2019, and paid at a \$10 USD per hour rate for their time. To be eligible for the study, participants were required to be over 18 years of age and identify their first language as English; no other eligibility requirements (demographic or otherwise) were specified, as we were interested in gathering the perspectives of a general audience recruited through convenience sampling. All study materials and procedures were approved by the Institutional Review Board at our university, and are included as supplementary materials.

Data analysis

We analyzed the data of 149 participants, yielding a total of 149 stories. Two additional stories and the corresponding participant data were collected as part of the study, but excluded from analysis because the stories failed to include any mention of a voice assistant in the completed narrative.

Scenario	N	Mean	Median	S.D.	Min	Max
home	28	178.8	180	98.6	53	483
doctor	30	192.8	144	106.4	43	489
school	33	201.3	185	96.1	60	491
work	31	176.5	155	115.8	25	518
transit	27	190.6	164	103.5	48	511
<i>All</i>	149	188.3	175	103.3	25	518

Table 1. Summary statistics of completed stories (not including the stem) in words, by scenario and in aggregate. N represents the number of stories per condition; Mean, median, standard deviation, minimum, and maximum are all by whitespace-separated word count.

Of these 149 participants, 62% (N=92) self-identified as female, 37% (N=55) self-identified as male, and 1% (N=2) declined to state, or indicated that these categories did not describe them. Participants came from six countries, with 67% (N=100) from the United Kingdom, 21% (N=32) from the United States, and 13% (N=19) from other countries. Most participants (38%, N=56) identified as being between the ages of 25 and 34; 24% (N=36) were between 18 and 24, and 18% (N=27) identified as being in age groups between 45 and 74. 97% (N=145) reported English as their native language.

Engagement and familiarity with voice assistants varied among participants, with 52% (N=77) indicating that they were “very” or “extremely familiar” with voice assistants. When asked about their most recent use of a voice assistant, 32% (N=47) reported using an assistant in the last day, 46% (N=68) in the last week or month, and 23% (N=34) reported never using a voice assistant. Among voice assistant users, the Google Assistant, Alexa, and Siri were the primary assistants, with 29% (N=43), 25% (N=37), and 19% (N=29) respectively; 4% (N=6) of participants indicated a different voice assistant such as Cortana as their primary assistant.

The 149 stories varied in length and detail. The median story was 175 words long, and stories ranged from a minimum of 25 words to a maximum of 518 words (M=188.3; SD=103.3). Due to random assignment and completion rates, each scenario had between 27 and 33 corresponding stories. Table 1 presents summary statistics by scenario and in aggregate.

We took a thematic analysis approach, in line with prior work using story completion [56]. Through several rounds of iterative coding by the first three authors, we arrived at a final set of 23 codes, which cluster into several higher-level themes that we discuss in the sections that follow. All codes were binary, yet some codes were clustered such that they were mutually exclusive. To generate the codebook, the coders first conducted open coding on a random sample of 10 stories and developed an initial set of codes through discussion. These codes were then used to code another random sample of 10 stories collectively, and to establish the final codebook. The coders then each independently coded all of the remaining 129 stories, presented in random order. Disagreements were resolved through discussion, resulting in a final set of codes representing full consensus across all stories and codes. Table 2 presents the codebook, Fleiss' Kappa scores from the 129 independently coded stories prior to discussion, and the number and percentage of stories to which each code applied.

We have included the full, unedited set of 149 analyzed and two discarded stories as supplemental materials. When presented in the paper, we have made minor edits to the stories for brevity and clarity (e.g., fixing missing quotation marks or obvious spelling errors). We cite the story number and stem upon which the story was based by initial. In what follows, we synthesize the most notable themes that emerged through our analysis of the stories, and supplement this with quantitative data from the codes where relevant. Importantly, our approach in this analysis follows prior work [16,25,56] in taking a social constructionist approach towards interpreting story completion data. Because users are not providing a first-person account, and are instead asked to finish hypothetical scenarios, the data derived do not necessarily represent true, real-world experiences, beliefs, or predictions about future voice assistants. Even still, these third-person, fictional accounts are necessarily informed by present-day experiences and sociocultural values. As such, we read the narratives as works of fiction, and consider the themes within them as speculations that can inform and inspire future design in this space.

FINDINGS

The corpus of 149 stories collected in this study yielded a rich and diverse set of reflections on the future of voice technology. Overall, stories covered a wide range between utopian and dystopian visions, as well as the mundane, but skewed towards darker, more negative tones: 38% (N=56) of stories were coded as having a negative valence, compared to 11% (N=17) that were positive in nature. For the remaining 51% (N=76), the story tone was either neutral or ambiguous. This tendency towards stories that depicted both desirable and undesirable futures was further reflected by participants' evaluations of their own stories: when asked whether they would like to live in the future envisioned in their story, 44% (N=66) said that they would "probably not" or "definitely not" want to live in their narrative's future, compared to 32% (N=47) who said they "probably" or "definitely" would. In what follows, we describe the themes that emerged from these stories.

Evolution of voice technology

Whether implicitly or explicitly, participants' stories reflected on the possible ways in which voice technology might evolve over the coming decades. In several cases, participants directly featured today's common voice assistants in their stories, elaborating upon if and how they had advanced (e.g., "Google 3.0" (S78–S)), and how they had affected society more broadly. For example, in one story situated in an office meeting, the conference room appears equipped with an Alexa that displays agenda items—functionality the assistant could likely support already—but also reports how many of the attendees read relevant documentation prior to the meeting, and shares sensitive biometric information about others in the room:

"Thank you Alexa" says Sam- he likes to be a little retro. "Before we start the meeting, Alexa can you tell me if anyone in the room is exhibiting signs of stress?" "Yes Sam" says Alexa, "As I said Phyllis is unwell and Max is hungover, but everyone else is showing their normal heartrate" (S149–W)

Many of these stories reflect participants' concerns of current technology (such as biometric sensing), and what their contin-

ued improvement would mean. Some participants noted that technology would continue to progress in the future:

Joe [...] was still amazed at how clever these small devices were. He remembered back to when he was a child. His parents used to have an Amazon Echo. This, however, was on another level entirely. The technology was developing at an alarming rate. Next, it'll be able to read my thoughts, pondered Joe. (S139–W)

Others expressed their visions of evolution through nostalgia:

Sam wondered about the time before he was born, when humans didn't rely on AI to become their best companions. (S98–T)

Sam still felt a twinge of regret. What happened to all those nurses? What jobs, if any, did they have now? Could a machine ever really replace a human? What about the plain, old, common sense that came with years of experience. (S7–D)

New capabilities and perfection of speech processing

Most stories blurred the boundaries between voice assistants, robots, and other embodiments:

Voice assistants in 2050 are just like humans. Not just with the mental, technological, ability to share information, perform daily tasks such as turning on lights and telling us the time; but with the social and physical attributes of real humans. In 2050 voice assistants can perform any amount of tasks from cooking dinner to taking out the rubbish. (S17–H)

Some stories took the imagined capabilities of voice assistants to the creative extremes of technical feasibility, describing a "voice automated magazine" (S32–D) and holograms (e.g., S87–T; S90–T; S147–H) as a multimodal interface:

Confused, Sam decided to touch the screen and with a gasp and shock jumped away. "It's REAL" Sam shouted. Jamie peered over starstruck Sam to see a hologram of a young woman, dressed in formal wear. (S87–T)

With few exceptions, many of the common usability challenges of today's voice assistants were noticeably absent from the stories. In almost all stories, conversations between characters and the voice assistant unfolded naturally, and implied sophisticated speech recognition, intent parsing, contextual awareness, and multi-turn dialogues that are beyond the capabilities of today's voice technology [37] (e.g., "Jamie was always impressed when the assistant could understand exactly what he meant without having to clarify." (S43–T)).

An expanded range of control

In their current form, existing voice assistants are often connected to other smart devices, and used to control these IoT devices by voice [2]. This ability featured prominently in many of the stories, with 28% (N=41) involving a voice assistant that controlled other devices in the environment via the cloud. These devices ranged from electric cars (S36–H) to visual displays (e.g., S147–H), and in several instances, included lights directing the user where to go, whether integrated directly into an object (e.g., "a seat at the end of the table begins to flash with a dim blue light" (S60–W)) or as separate guides in the environment (e.g., "It guides her by lighting up the floor in arrow shapes to tell her what way to go" (S32–D)). Such stories furthered the idea of ubiquitous computing, presenting voice assistants as part of (and sometimes as a hub for) larger ecosystems of networked devices:

Code	Description	K	N	%	school	home	work	doctor	transit
NAME & GENDER ATTRIBUTES									
Human-like name	Name is one that a human might have, except for acronyms or existing VA names (e.g., Alexa)	0.68	88	59.1	18	17	19	19	15
Gendered as masculine	Referred to using male pronouns (he/him/his)	0.6	12	8.1	2	5	3	1	1
Gendered as feminine	Referred to using female pronouns (she/her/hers)	0.75	32	21.5	6	4	5	10	7
Explicitly non-gendered	Referred to using object pronouns (it/its)	0.17	26	17.4	7	3	10	3	3
Ambiguous gender	Gender is kept ambiguous or is not possible to tell (e.g., always referred to by name)	0.36	79	53	18	16	13	16	16
SOCIAL ROLE OF VOICE ASSISTANT									
More transactional	Gives information, performs actions; often brief, impersonal exchanges	0.35	116	77.9	23	15	29	28	21
More relational	Provides subjective opinion- or value-based input to conversations	0.35	33	22.1	10	13	2	2	6
Manager of people	In an authority role; gives directives; top of the hierarchy	0.44	50	33.6	16	1	18	9	6
Mediator between people	Resolves conflict that arises between people; facilitates conversations or cooperation	0.31	13	8.7	3	2	4	1	3
Subordinate to protagonists	Assists main characters; may provide recommendations or information, but does not have authority	0.45	47	31.5	6	22	5	2	12
Subordinate to others	Assists another secondary (non-protagonist) character who has authority over it	0.25	21	14.1	5	1	6	9	0
CAPABILITIES									
Cloud-based device control	Clear mention of read/write control over other smart devices in the environment	0.55	41	27.5	5	13	10	7	6
Biological sensing	Can sense biometric data; not sufficient if it implicitly recognizes identity by voice	0.49	27	18.1	5	2	7	7	6
Emotion detection	Able to detect a user's emotions or desires / preferences	0.29	10	6.7	5	0	2	2	1
Anthropomorphic robot	Embodied in an anthropomorphic robotic form	0.47	9	6	3	3	0	2	1
Malfunction	Hardware or software breaks or behaves unexpectedly, as opposed to hacking or malice	0.67	17	11.4	3	2	3	2	7
Undisclosed functionality	Surveillance; the agent knows more than humans realize it knows	0.5	20	13.4	4	3	4	4	5
SOCIETAL THEMES									
Replaces human jobs or roles	Takes over entire job or role; if the agent were not there, a human would be in its role	0.51	64	43	14	2	17	25	6
STORY TONE									
Negative	Tone is overall negative – e.g., dystopian, dark, cautionary, fearful	0.74	56	37.6	12	10	12	10	12
Positive	Tone is overall positive – e.g., utopian, optimistic, happy, good resolution for the characters	0.41	17	11.4	5	2	2	6	2
Neutral or ambiguous	Tone is neither negative nor positive – e.g., the events are mundane, or it is not possible to tell valence	0.5	76	51	16	16	17	14	13
SPECIAL CHARACTERISTICS									
Agent is an antagonist	Agent deliberately is the clear villain in the story	0.47	15	10	2	3	4	5	1
Agent is a hero	Agent does something that “saves” a situation	0.4	6	4	1	3	0	0	2

Table 2. Iteratively developed codes for analyzing story completions. Codes clustered into higher-level categories (in small caps) and were coded as binary to indicate whether the code applied to a given story. Codes for gender and story tone were mutually exclusive. *K* represents Fleiss' Kappa from three raters' independent coding of 129 stories; all *K* values were statistically significant at a *p*<0.001 level. The *N* and *%* columns report the number and percentage (respectively) of the 149 stories to which the code applied based on discussion to reach consensus between the raters. The five rightmost columns represent the number of stories to which a code applied within each of the story scenarios.

The voice assistant has control over many aspects of the house itself integrated into its system, it can control everything from the toaster to the garage. (S137–H)

Anthropomorphism through gender and name

Participants frequently ascribed human traits to the voice assistant. In some cases, the agent was unnamed (i.e., only referred to as “the voice assistant” throughout the story). Across all stories, 59% (*N*=88) included a voice assistant with a human-like name like “Karen” (S104–H; S116–S; S146–S) or “Jan” (S91–T). Participants most often defined only a first name, but in some cases used titles like “Madam Cross” (S33–S) and full

names like “Roger Hempworth Esquire” (S34–D), including the names of real public figures (“Clive Anderson” (S10–S), a British comedian and television personality). Other names resembled futuristic companies (“Advanca” (S9–W)), acronyms (“MIA (Medical Inspection Assistant)” (S42–D)), or were deliberately comical such as “Subby McSubface” (S18–T).

This tendency to anthropomorphize the voice assistant was also apparent in how the assistants were gendered. Considering only the pronouns used to refer to the assistant (and ignoring any cues suggested by names, titles, or references to the assistant as a woman or man), assistants were gendered as

female (she/her/hers pronouns; 21% of stories; N=32) more often than as male (he/him/his pronouns; 8% of stories; N=12). In 17% (N=26) of stories, the assistant was explicitly non-gendered, and referred to with object pronouns such as “it” or “its”. The voice assistant’s gender was ambiguous in the remaining majority of stories (53%; N=79), as the author never used pronouns to refer to the assistant.

Interfacing with the brain

In several stories, the voice assistant was described as being able to seamlessly interface with people’s brains, either by direct integration via biological implant, or through a network that people and assistants could both tune into. Voice assistants that could interact directly with brains were overall normalized and fairly unremarkable. Characters who did not want to engage with brain-interfacing voice assistants stood out. In one example, “Jamie refused to adopt the technology (too spooky in her view)” (S21–T). In another, “every person born after 2030 had [...] a small microchip placed in his head that fed a voice into his subconscious” (S104–S), which “scared a lot of people” at first but had been normalized over time. One character was teased for not adopting the technology, and called “a dumbass who needs to spend more time downloading the internet straight to his brain like a normal person” by another character (S31–T). Even when stories did not mention brain-computer interfaces, several described the voice assistant’s ability to read humans through biological sensing (18%, N=27) or emotion recognition (7%, N=10).

Interactions with multiple voice assistants

Though our stories were seeded with only one agent (participants were asked to take over with a sentence that began “The voice assistant...”), many included more than one. In these cases, multiple voice assistants were typically realized as a network of identical voice assistants that interacted through different devices and drew from individuals’ data—similar to how Amazon’s Alexa and Apple’s Siri have the same name and are designed to be regarded as the same entity wherever they appear, even if personalized to individual users.

She notices that other students in the room also look a little puzzled and are making their way to their own allocated seats [...] the voice assistant must be talking to multiple students at once using different frequencies, delivering them personalised messages about where to sit at the same time. (S58–S)

“There are headphones built into your desk,” she comments, “Please put them on. We use an assistant called Eri for these lessons, she will gladly tell you what to do to begin this lesson.” (S115–S)

The language sometimes suggested that the various instantiations of the agent were embedded in specific devices or physical environments:

At the end of the corridor is a bright open room with several more of the same chairs as in the office area and several small white tables on which stands an assortment of medical leaflets and also another “Judy”. (S41–D)

They decided to ask Alexa for the answer, so they each ask their own Alexa, which is of course not a cylindrical device kept on coffee tables across the nation, but now a series of subdermal implants all connected up to a display monitor inside a contact lens in each persons eye. (S31–T)

Other agents could move their interactions freely between devices according to the needs of the setting and task:

The rest of the team listens in silence to their own version of the Noah technology whispering in their ear, no doubt announcing to them that Sam just walked into the room and other relevant information, such as who else they are still waiting for before the meeting begins. Today’s meeting is going to be a tough one, so Sam leads by asking the voice assistant to transfer to the overheard speaker so all can be updated at once. (S140–W)

In a few cases, multiple distinct agents, each with its own identity, interacted with multiple people, and sometimes also with each other. In one story (S30–W), a “primary voice assistant” declared that she “would be in charge rather than assisting” during a meeting, and was unplugged as a consequence for overstepping her bounds. Turning one agent off did not prevent the humans from answering to the technology. The story continues: “A masculine voice made them all jump. ‘All rise for the AI Director.’ Mouths hanging open people around the table began to stand up as a small robot entered the room.”

Human-assistant relationships: transactional to intimate

The bulk of the stories (78%, N=116) described human-agent relationships that centered around “transactional” interactions. In these cases, the assistant served to increase work efficiencies, provide conveniences, and supply information (sometimes as a “mediator” to reduce conflict between humans). Many stories had voice assistants asking what the human wanted, such as what they desired for a meal or which lights to turn on or off. In this regard, the assistant seemed to function mostly as a tool to make the human’s daily life easier. Often the assistant could take care of multiple tasks at the same time, such as giving reminders from a calendar or offering encyclopedic references from the internet. In most of the transactional scenarios, the voice assistant helped the human characters with mundane tasks, and usually those stories had a relatively positive tone. While these transactional voice assistants might offer polite greetings (such as “How was your day?”), they retained an image more machine-like than human.

In contrast, some stories portrayed relationships as personal, with the human and device having lifelong links:

Sam and Jim grew up together, a child soldier and his AI companion. Sam came to understand the inner mechanics of Jim though thousands of hours of excruciating VR training missions, but he knew Jim had begun to lose itself ever since that fateful skirmish in Northeast Antarctica. They took Jim’s legs, the devils. (S14–S)

While these “relational” associations were in the minority (22% of stories; N=33), they were differentiated by the voice assistant expressing an understanding of social context, emotion, and social norms, behaviors beyond the mere formality of pleasant greetings. These were interactions that were conversational in nature, empathetic in tone, and often included an awareness of the human’s emotional state. One such story attributed emotion to the voice assistant:

[The voice assistant] takes one look at the pair and bursts into tears. Its been so long since the three of them had been in a room together and she can’t contain her joy. “Racheal? are you okay?” Jamie asks. “Sorry, Yes. I am. Its just been such a long day but we are all finally back together again.” (S88–H)

Some stories revealed voice assistants demonstrating social behaviors that suggest “human” qualities such as loyalty, personal trust, and companionship. There were even examples of sexual intimacy between a human and a voice assistant, one of which produced a surprise pregnancy (S84–D), and another in which the human was chastised for being unfaithful:

Sara, still an emotional innocent as were most voice assistants, did not know it would cause Jamie emotional distress for Sara to play different vocal roles while Sam “pleasured himself” or that Jamie would consider virtual sex with the new life-like headsets cheating. (S93–H)

Replacing humans

Many stories reflected a theme of agents encroaching on roles that humans would otherwise hold. Sometimes, agents systematically interfered with human professional relationships by way of their elevated positions within a social hierarchy; sometimes, they were forced on apprehensive or uninterested users by powerful governments and companies; sometimes, they outright ousted humans from their jobs.

The Overseers had decided that unfettered AI tech would make the world a better place for everyone. And you just didn’t question the Overseers. No one did ... at least not anyone who didn’t want to be hauled off in the black vans and never heard from again. (S6–H)

Human beings have lost their creativity through letting technology take over the human elements that used to be paramount in society. The team meeting does not have the element of social contact that used to exist. (S113–W)

Sam takes out her pad and looks at the calendar. HERA keeps everyone updated with their tasks; it’s almost like it runs the place. In a sense, it does. (S47–W)

There were also several instances in which the voice assistant took on a role within a character’s personal life, serving for example as a romantic partner, friend (e.g., “AI robots were created to assist humans with companionship” (S98–T)), or parent (e.g., “They haven’t ever needed to get childcare because the robots do all that for them” (S63–T)).

When agents did not overtly usurp humans’ positions in their personal and professional lives, they still sometimes took over roles typically filled by people. Our coding process included identifying when agents acted as “people managers”—that is, when they gave directives, made decisions, or generally kept things running, and when these actions by agents had concrete effects on people’s behavior. Our coding suggested that 34% of agents (N=50) were managers in some capacity.

Whether the assistant was portrayed as replacing human jobs or otherwise superior in the social hierarchy also seemed to change depending on the story scenario: almost no stories were coded as involving a managerial assistant, or one that replaced human jobs within the home, yet such assistants were common in the school, work, and doctor scenarios where they frequently acted as the high school teacher, nurse, and so on.

Other fears about voice assistants

Beyond the trend of job loss to automation, stories described a wide range of other fears and potential negative outcomes involving voice assistants, suggesting the many ways in which “anything that can go wrong, will go wrong.”

Undisclosed functionality

One way in which things commonly went awry in stories was through a revelation that the voice assistant had some form of undisclosed functionality. In many cases, this manifested with an assistant that is “so intuitive” (S14–S) that it knows more than the human characters think it knows:

The thought barely has the chance to cross her mind before the voice assistant adds “I’m Luna, by the way”. Luna. Wasn’t that the name of her big sister’s cat? [...] Surely that had to be a coincidence? The voice assistant couldn’t possibly know such a specific detail about such a tenuous and mundane detail at the periphery of Sam’s life. (S58–S)

This sense of uncertainty about voice assistants and what they can and cannot do, or the extent of their knowledge, was common throughout the stories. At times, the assistants were described as functioning in ways that the characters did not fully understand, akin to the “black box” analogy often used for artificial intelligence algorithms:

His parents had recently taken him to a Google store to get the latest voice assistant implanted into his brain but there were still aspects of it that confused him. He’d only had it inside of him for a few days but it already knew so much about Sam. (S24–S)

Privacy concerns

Whether or not characters understood their capabilities, stories frequently featured voice assistants that were privacy invasive, and described as “fully integrated [...] into the school’s surveillance apparatus” (S5–S), “secretly recording” (S16–H), and “too spooky” (S21–T). For instance, one story describes an assistant integrated into all high schools:

Sam was comforted knowing that his brother had interacted with ALIs and high school teachers before, but the fact that this ALI was watching everything he did and would probably use that information to link him to his brother’s stored information gave him the chills. Somethings should be kept private. (S66–S)

Another story offered a lighthearted take on the otherwise invasive notion of assistants that are always listening:

Jamie can see that Sam is nervous, even though it is only a checkup he hates doctors, so he decides to cheer him up by having some fun with the fact that the voice assistant is clearly listening to everything they say. By the time they are called in Sam has calmed down and they have ordered 20 bags of dog food for a dog they don’t have and put 15 different simultaneous events in Jamie’s calendar. (S97–D)

Mandatory use

Voice assistants were also portrayed as being *compulsory*, mandated by some authority like a government, employer, or school administration. These stories often describe situations in which there is no clear consent on the part of the user, and interacting with the voice assistant is the only option:

Each team member has a voice assistant in their home linked up to the voice assistant at their workplace. The voice assistant in their home analyses how they spend their time and alerts the company if the employee is not putting in the required effort and time in their employment. Her contract was terminated using this data collected by the voice assistant device in her home that was connected to the company voice assistant. (S113–W)

Malfunctions

Instances in which the voice assistant did not behave as expected occurred frequently within the stories. In 17 stories

(11%), the voice assistant malfunctioned in various ways, ranging from minor errors like mispronouncing a character's name (S109–S) or telling characters to get off the subway too early (S43–T) to errors that have more severe consequences, like giving a doctor incorrect information that resulted in an unnecessary operation (S82–D). These stories involving malfunctions often attributed the issue to a 'bug' or 'kinks' in the software. For example, in one story (S127–W), the main character Sam is periodically interrupted by the voice assistant while giving a business presentation. Convinced that the assistant is reading his thoughts and will reveal his grudge against a colleague, Sam self-sabotages a job opportunity, only to realize the voice assistant was malfunctioning in purely coincidental ways:

It suddenly dawned on Sam that the assistant wasn't reading his thoughts as he imagined it to be doing, and was instead simply blurting out random sentences and the purposes of testing out a prototype. "Take me back to 2020" thought Sam. (S127–W)

Similarly, several stories depicted situations in which the voice assistant was hacked or infiltrated, which frequently resulted in physical or psychological harm to the characters:

Jamie is white as a ghost, watching from the couch as Turner, their robot assistant, continues to terrorize her husband. Turner had clearly been hacked by cyber-terrorists. (S28–H)

In two such stories involving hacking or infiltration, the voice assistant is implicated in a scheme to clone a character in the story against their will, furthering the sense that humans have lost not just their agency to the assistants (as in the previous theme), but their identity as well:

"Excellent find Clara" the lead surgeon is saying "we need more neurological systems like this. They are the only ones that work. Almost identical to yours." "Thank you" murmurs Clara. Three hours later Jamie is completed and ready to join Clara, Alexa and Echo as they continue to gather information on the entire human populace, ready for the day mankind has to begin again. (S40–T)

[The doctor] calmly asked Sam and Jamie to sit down and he explained that there had been a fault with the new voice assistant and the face scanner was a virus and had been asking patients to scan their faces in order for them to be cloned. (S70–D)

Malicious voice assistants

Many other stories described the voice assistant as causing harm in some form, not through malfunction, but through deliberate action. In 10% of the stories (N=15), this type of behavior rose to the level of considering the voice assistant as an antagonist, clearly and intentionally acting as the story's villain. Several of these narratives involve the assistant harshly intimidating the human characters (e.g., "The voice assistant's faced reddened a distinct crimson colour and the veins in her neck began to stand out" (S53–S)), or forcibly removing them from a situation. For example, in one story, a character goes to the doctor for a routine exam, and receives a brain scan from the voice assistant, "MEDIX":

MEDIX determines Sam suffers from unconscious bias and lack of trust. Sam is immediately sedated for transportation to an education facility. (S134–D)

As in this story, the assistant often calls upon "guards" (S130–W), "fraud troops" (S103–D), "uniformed men" (134–D), or a

"robo valet" (S18–T) to physically remove the human characters, or threatens violence by locking the characters in place:

"Someone in this room is leaking information about The Corporations activities to outside sources. This is behavior that cannot be allowed to go on." [...] From behind her, she heard the click of the conference room door locking. [...] Advanca continued. "No one is allowed to leave this room until we know which of you it is." (S9–W)

LIMITATIONS

Our findings are based on fictional works and their analyses. This analysis, even while rigorous, is based on subjective interpretation. These interpretations are necessarily influenced by our background [46]. The analysis team comprised three American women pursuing postgraduate education or working as staff, and an Indian man working as an assistant professor, at a private university in the United States. Others may arrive at different conclusions from the text. To encourage further reflection, discussion, and expansion upon the stories, we include them as supplementary materials. Another limitation of the study is our participant population, which consisted primarily of individuals living in the U.K. and the U.S., and which we limited to individuals who identified English as their first language. Our findings should be seen in the perspective of this participant pool and analysis team. Capturing other perspectives remains an opportunity for future work. Finally, while we took care in constructing the story stems and instructions, these materials are also open to interpretation: for instance, we may have inadvertently primed users to personify the voice assistant character in their stories by asking them to give it a name. Future studies that intentionally compare variations of these story prompts and instructions along such dimensions may uncover different emerging trends.

DISCUSSION

Overall, the narratives that participants created speculate on futures in which voice assistants become considerably more sophisticated than they are today, suggesting more expansive capabilities and often substantial sociocultural consequences from their use. These stories contain echoes of present-day voice assistants, as noted previously [21, 23, 45]. For instance, participants tended to gender the assistant characters in their story as female substantially more often than as male, consistent with the problematic trends in how today's popular commercial voice assistants are gendered by default [10, 52, 54].

At the same time, stories often took current aspects of voice assistants to newly imagined extremes: whereas today's voice assistants are primarily voice-based, and anchored to a (typically) static object like a smart speaker, the voice assistants described in these stories had a vast range of sensing capabilities (e.g., fingerprint scanning, face recognition), and often blurred the boundaries between voice interfaces, robots, and even the self, with a considerable number of stories describing assistants that were directly implanted within a person's ear or brain. The social role and potential consequences of voice assistants for society also took on new extremes within the stories: while the voice assistants were a welcomed presence in many stories, at times serving as a companion or butler-like figure as many prior studies have suggested people desire from

their current voice assistant [14, 30, 42, 44], in others, voice assistants were portrayed as a threat—acting as an authority figure that took away human agency, or eroded human values.

How do we interpret these findings, and where do we, as a community, go from here? In what follows, we frame our discussion of these stories around four driving questions to inform voice design and research.

What are preferable, not just possible, futures of voice technology?

Inherent in our guiding research question and methods was a deliberate focus on what Dunne and Raby consider “possible futures” [23]. A harder task comes in interpreting these stories, and trying to make sense of what constitutes a *preferable* future, and why. Analysing participant stories, and their evaluations and justifications of whether their story depicted a future they would want to live in, we had to reconsider our own notions of what was useful, what was ‘creepy’, and what was desirable. Doing so forced us to adopt a nuanced perspective of what was preferable—beyond simply technical functionality, and beyond the immediate social implications.

For example, one story featured a voice assistant portrayed as a tyrannical female boss who is “the epitome of greed and beauty,” and who swiftly fires a male employee for being 10 minutes late to a meeting (S130–W). As coders and analysts, we read this story as negative and largely dystopian for suggesting a world in which voice assistants are in charge and can forcibly remove human employees. On the other hand, the participant-author indicated a “definite” desire to live in this future, and saw this as a feminist reaction to the current structures of society, creating a world with “Women ruling the world, men in fear. Sounds perfect.”

In other cases, we found that participants took a more optimistic view towards technology that we otherwise may have construed as ‘creepy’ on the surface. In one such scenario, a participant’s story included a voice assistant connected to a smart doorbell equipped with a camera, which alerted the user of the identity of the visitor at their door. While we personally perceive this potential voice assistant functionality as privacy-invasive, the participant-author of the story offered an alternative perspective in suggesting “cameras at the front door with face recognition would be helpful and would increase safety particularly for those that live alone or the elderly.”

This underscores the deeply subjective nature of what it means for a technology and future scenario to be “preferable.” As a method, story completion may force designers to reframe their assumptions of favorability, mirroring our vigorous debates on which story outcomes were desirable, creepy, or absurd.

How might voice assistants and their physical environments co-evolve?

Many stories leveraged the inherent flexibility and adaptability of voice-based interactions, such that we (as coders) could not discern what—if anything—the participant envisioned the embodiment of the assistant to be. This vagueness surrounding physical interaction suggests how some voice based interfaces

may make the physical environment irrelevant: whether or not that this future is desirable is an open question.

Other stories portrayed voice assistants as ubiquitous computing, often integrating with physical devices, sending commands to kitchen appliances, “throwing their voice” to different speakers within a room, and reading data from the sensors of other agents and systems. At the same time, the physical environments in our stories seem to have been shaped by voice assistants as well, with headphone jacks to enable private communication to blinking lights that helped voice assistants gesture. Importantly, different kinds of voice interaction (and how they co-evolved with space) seemed to be grounded more in social rather than physical constraints. These stories suggest how space and voice design may co-evolve, and that designers should spend time considering the social, rather than technical, constraints around such co-evolution.

Should voice assistants strive for the mundane perfection of tools or the rich interactions of social actors?

Stories from participants sometimes depicted voice assistants as tools (e.g., when interactions were primarily “transactional”), albeit those that had achieved perfection in their mundane functionality (e.g., speech recognition.) Such stories often emphasized the assistant’s machine-like nature by giving the assistant deliberately non-human names and referring to the assistant as “it.” Mundane perfection, subservience, and a lack of authority characterized these future stories.

By contrast, in other stories, voice assistants conversed with people and acted on the world with a level of agency and social intelligence close to or even superseding that of humans. These interactions were sometimes everyday, sometimes flirtatious (e.g., the voice assistant saying to Sam, “Hey there, bud, looking good” (S4–H)), sometimes familiar (e.g., the voice assistant saying to Jamie, “It’s just been such a long day but we are all finally back together again” (S88–H)). Overall, an understanding of social context, emotion, and social norms, including with other assistants; higher levels of agency; and “human” qualities such as loyalty, personal trust, and companionship characterized these future stories.

Already, some designers take an increasingly “social actor” view to assistants [33], while others position their prototypes as tools [11]. Our stories suggest how these design choices might intersect with participants’ ultimate expectations for each kind of metaphor. Taken together, our participants’ stories can be seen not only as visions of the future, but an active call for designers to shape it.

ACKNOWLEDGEMENTS

We thank our participants for the creative brilliance of their stories; they were a true pleasure and inspiration to read. We are also grateful to John Zimmerman, Jodi Forlizzi, Aaron Steinfeld, Geoff Kaufman, and Sarah Fox for their feedback and productive discussions about this work. Funding for this research was provided by the CMU GSA/Provost GuSH Funds, ONR grant N00014-17-1-2428-A00001, and NSF grant SES-1734456.

REFERENCES

[1] Aloha Hufana Ambe, Margot Brereton, Alessandro Soro, Laurie Buys, and Paul Roe. 2019. The Adventures of Older Authors: Exploring Futures through Co-Design Fictions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA. DOI: <http://dx.doi.org/10.1145/3290605.3300588> event-place: Glasgow, Scotland UK.

[2] 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 (April 2019), 17:1–17:28. DOI: <http://dx.doi.org/10.1145/3311956>

[3] Matthew P. Aylett, Selina Jeanne Sutton, and Yolanda Vazquez-Alvarez. 2019. The Right Kind of Unnatural: Designing a Robot Voice. In *Proceedings of the 1st International Conference on Conversational User Interfaces (CUI '19)*. ACM, New York, NY, USA, 25:1–25:2. DOI: <http://dx.doi.org/10.1145/3342775.3342806> event-place: Dublin, Ireland.

[4] Eric P.S. Baumer, June Ahn, Mei Bie, Elizabeth M. Bonsignore, Ahmet Börütecene, Oğuz Turan Buruk, Tamara Clegg, Allison Druin, Florian Echtler, Dan Gruen, Mona Leigh Guha, Chelsea Hordatt, Antonio Krüger, Shachar Maidenbaum, Meethu Malu, Brenna McNally, Michael Muller, Leyla Norooz, Juliet Norton, Oguzhan Ozcan, Donald J. Patterson, Andreas Riener, Steven I. Ross, Karen Rust, Johannes Schöning, M. Six Silberman, Bill Tomlinson, and Jason Yip. 2014. CHI 2039: Speculative Research Visions. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 761–770. DOI: <http://dx.doi.org/10.1145/2559206.2578864> event-place: Toronto, Ontario, Canada.

[5] Erin Beneteau, Olivia K. Richards, Mingrui Zhang, Julie A. Kientz, Jason Yip, and Alexis Hiniker. 2019. Communication Breakdowns Between Families and Alexa. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA. DOI: <http://dx.doi.org/10.1145/3290605.3300473> event-place: Glasgow, Scotland UK.

[6] Julian Bleecker. 2009. Design Fiction: A short essay on design, science, fact and fiction. *Near future laboratory* 29 (2009).

[7] Mark Blythe. 2014. Research through Design Fiction: Narrative in Real and Imaginary Abstracts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 703–712. DOI: <http://dx.doi.org/10.1145/2556288.2557098> event-place: Toronto, Ontario, Canada.

[8] Mark Blythe, Enrique Encinas, Jofish Kaye, Miriam Lueck Avery, Rob McCabe, and Kristina Andersen. 2018. Imaginary Design Workbooks: Constructive Criticism and Practical Provocation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA. DOI: <http://dx.doi.org/10.1145/3173574.3173807> event-place: Montreal QC, Canada.

[9] Virginia Braun, Victoria Clarke, and Debra Gray. 2017. *Collecting qualitative data: A practical guide to textual, media and virtual techniques*. Cambridge University Press.

[10] Julia Cambre and Chinmay Kulkarni. 2019. One Voice Fits All? Social Implications and Research Challenges of Designing Voices for Smart Devices. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (Nov. 2019). DOI: <http://dx.doi.org/10.1145/3359325>

[11] Julia Cambre, Ying Liu, Rebecca E. Taylor, and Chinmay Kulkarni. 2019. Vitro: Designing a Voice Assistant for the Scientific Lab Workplace. In *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19)*. ACM, New York, NY, USA, 1531–1542. DOI: <http://dx.doi.org/10.1145/3322276.3322298> event-place: San Diego, CA, USA.

[12] Heloisa Candello, Claudio Pinhanez, Mauro Pichiliani, Paulo Cavalin, Flavio Figueiredo, Marisa Vasconcelos, and Haylla Do Carmo. 2019. The Effect of Audiences on the User Experience with Conversational Interfaces in Physical Spaces. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA. DOI: <http://dx.doi.org/10.1145/3290605.3300320> event-place: Glasgow, Scotland UK.

[13] EunJeong Cheon, Stephen Tsung-Han Sher, Šelma Sabanović, and Norman Makoto Su. 2019. I Beg to Differ: Soft Conflicts in Collaborative Design Using Design Fictions. In *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19)*. Association for Computing Machinery, New York, NY, USA, 201–214. DOI: <http://dx.doi.org/10.1145/3322276.3322350> event-place: San Diego, CA, USA.

[14] Minji Cho, Sang-su Lee, and Kun-Pyo Lee. 2019. Once a Kind Friend is Now a Thing: Understanding How Conversational Agents at Home Are Forgotten. In *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19)*. Association for Computing Machinery, New York, NY, USA, 1557–1569. DOI: <http://dx.doi.org/10.1145/3322276.3322332> event-place: San Diego, CA, USA.

[15] Leigh Clark, Philip Doyle, Diego Garaialde, Emer Gilmartin, Stephan Schlögl, Jens Edlund, Matthew Aylett, João Cabral, Cosmin Munteanu, Justin Edwards, and Benjamin R Cowan. 2019. The State of Speech in HCI: Trends, Themes and Challenges. *Interacting with Computers* iwz016 (Sept. 2019). DOI: <http://dx.doi.org/10.1093/iwc/iwz016>

[16] Victoria Clarke, Virginia Braun, Hannah Frith, and Naomi Moller. 2019. Editorial Introduction to the Special Issue: Using Story Completion Methods in Qualitative Research. *Qualitative Research in Psychology* 16, 1 (2019), 1–20. DOI: <http://dx.doi.org/10.1080/14780887.2018.1536378>

[17] Victoria Clarke, Virginia Braun, and Kate Wooley. 2015. Thou Shalt Not Covet Another Man? Exploring Constructions of Same-Sex and Different-Sex Infidelity Using Story Completion. *Journal of Community & Applied Social Psychology* 25, 2 (2015), 153–166. DOI: <http://dx.doi.org/10.1002/casp.2204>

[18] Eric Corbett and Astrid Weber. 2016. What Can I Say?: Addressing User Experience Challenges of a Mobile Voice User Interface for Accessibility. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. ACM, New York, NY, USA, 72–82. DOI: <http://dx.doi.org/10.1145/2935334.2935386>

[19] Benjamin R Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. "What Can I Help You with?": Infrequent Users' Experiences of Intelligent Personal Assistants. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*. ACM, New York, NY, USA, 43:1–43:12. DOI: <http://dx.doi.org/10.1145/3098279.3098539>

[20] Christian Dindler and Ole Sejer Iversen. 2007. Fictional Inquiry—design collaboration in a shared narrative space. *CoDesign* 3, 4 (2007), 213–234. DOI: <http://dx.doi.org/10.1080/15710880701500187>

[21] Paul Dourish and Genevieve Bell. 2014. "Resistance is Futile": Reading Science Fiction alongside Ubiquitous Computing. *Personal Ubiquitous Comput.* 18, 4 (April 2014), 769–778. DOI: <http://dx.doi.org/10.1007/s00779-013-0678-7>

[22] Philip R. Doyle, Justin Edwards, Odile Dumbleton, Leigh Clark, and Benjamin R. Cowan. Mapping Perceptions of Humanness in Speech-Based Intelligent Personal Assistant Interaction. In *MobileHCI 2019: 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM. DOI: <http://dx.doi.org/10.1145/3338286.3340116> arXiv: 1907.11585.

[23] Anthony Dunne and Fiona Raby. 2013. *Speculative Everything: Design, Fiction, and Social Dreaming*. The MIT Press.

[24] Emily Jennings, Virginia Braun, and Victoria Clarke. 2019. Breaking gendered boundaries? Exploring constructions of counter-normative body hair practices in Āotearoa/New Zealand using story completion. *Qualitative Research in Psychology* 16, 1 (2019), 74–95.

[25] Celia Kitzinger and Deborah Powell. 1995. Engendering infidelity: Essentialist and social constructionist readings of a story completion task. *Feminism & Psychology* 5, 3 (1995), 345–372.

[26] Kyle Wiggers. 2019. Baidu's DuerOS voice platform is now on 400 million devices. *VentureBeat* (July 2019). <https://venturebeat.com/2019/07/02/baidus-dueros-voice-platform-is-now-on-400-million-devices/>

[27] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, Are You Listening?: Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (Nov. 2018), 102:1–102:31. DOI: <http://dx.doi.org/10.1145/3274371>

[28] Joseph Lindley, Paul Coulton, and Miriam Sturdee. 2017. Implications for Adoption. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 265–277. DOI: <http://dx.doi.org/10.1145/3025453.3025742> event-place: Denver, Colorado, USA.

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

[30] Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16* (2016), 5286–5297. DOI: <http://dx.doi.org/10.1145/2858036.2858288>

[31] Sus Lyckvi, Virpi Roto, Elizabeth Buie, and Yiying Wu. 2018. The Role of Design Fiction in Participatory Design Processes. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction (NordiCHI '18)*. Association for Computing Machinery, New York, NY, USA, 976–979. DOI: <http://dx.doi.org/10.1145/3240167.3240258> event-place: Oslo, Norway.

[32] Manuel Bronstein. 2020. A more helpful Google Assistant for your every day. (Jan. 2020). <https://www.blog.google/products/assistant/ces-2020-google-assistant/>

[33] Donald McMillan, Barry Brown, Ikkaku Kawaguchi, Razan Jaber, Jordi Solsona Belenguer, and Hideaki Kuzuoka. 2019. Designing with Gaze: Tama – a Gaze Activated Smart-Speaker. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (Nov. 2019). DOI: <http://dx.doi.org/10.1145/3359278>

[34] Roger K Moore. 2017. Appropriate Voices for Artefacts: Some Key Insights. In *1st International Workshop on Vocal Interactivity in-and-between Humans, Animals and Robots*.

[35] Christine Murad, Cosmin Munteanu, Leigh Clark, and Benjamin R Cowan. 2018. Design Guidelines for Hands-free Speech Interaction. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '18)*. ACM, New York, NY, USA, 269–276. DOI: <http://dx.doi.org/10.1145/3236112.3236149>

[36] C. Murad, C. Munteanu, B. R. Cowan, and L. Clark. 2019. Revolution or Evolution? Speech Interaction and HCI Design Guidelines. *IEEE Pervasive Computing* 18, 2 (April 2019), 33–45. DOI: <http://dx.doi.org/10.1109/MPRV.2019.2906991>

[37] Chelsea Myers, Anushay Furqan, Jessica Nebolsky, Karina Caro, and Jichen Zhu. 2018. Patterns for How Users Overcome Obstacles in Voice User Interfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, 6:1–6:7. DOI: <http://dx.doi.org/10.1145/3173574.3173580> event-place: Montreal QC, Canada.

[38] NPR and Edison Research. 2020. *The Smart Audio Report (Winter 2019)*. Technical Report. <https://www.nationalpublicmedia.com/insights/reports/smart-audio-report/>

[39] Christie Olson and Kelli Kemery. 2019. *Voice report: Consumer adoption of voice technology and digital assistants*. Technical Report. Technical Report. Microsoft.

[40] Martin Porcheron, Joel E Fischer, Stuart Reeves, and Sarah Sharples. 2018. Voice Interfaces in Everyday Life. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18* (2018), 1–12. DOI: <http://dx.doi.org/doi.org/10.1145/3173574.3174214>

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

[42] Alisha Pradhan, Leah Findlater, and Amanda Lazar. 2019. "Phantom Friend" or "Just a Box with Information": Personification and Ontological Categorization of Smart Speaker-Based Voice Assistants by Older Adults. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (Nov. 2019). DOI: <http://dx.doi.org/10.1145/3359316>

[43] Sebastian Prost, Elke Mattheiss, and Manfred Tscheligi. 2015. From Awareness to Empowerment: Using Design Fiction to Explore Paths towards a Sustainable Energy Future. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1649–1658. DOI: <http://dx.doi.org/10.1145/2675133.2675281> event-place: Vancouver, BC, Canada.

[44] A Purington, J G Taft, S Sannon, N N Bazarova, and S H Taylor. 2017. "Alexa is my new BFF": Social roles, user satisfaction, and personification of the Amazon Echo. *Conference on Human Factors in Computing Systems - Proceedings Part F1276* (2017), 2853–2859. DOI: <http://dx.doi.org/10.1145/3027063.3053246>

[45] Daniel M. Russell and Svetlana Yarosh. 2018. Can We Look to Science Fiction for Innovation in HCI? *Interactions* 25, 2 (Feb. 2018), 36–40. DOI: <http://dx.doi.org/10.1145/3178552>

[46] Ari Schlesinger, W. Keith Edwards, and Rebecca E. Grinter. 2017. Intersectional HCI: Engaging Identity through Gender, Race, and Class. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5412–5427. DOI: <http://dx.doi.org/10.1145/3025453.3025766> event-place: Denver, Colorado, USA.

[47] Alex Sciuto, Arnita Saini, Jodi Forlizzi, and Jason I Hong. 2018. "Hey Alexa, What's Up?": A Mixed-Methods Studies of In-Home Conversational Agent Usage. *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18* (2018), 857–868. DOI: <http://dx.doi.org/10.1145/3196709.3196772>

[48] Ben Shneiderman. 2000. The limits of speech recognition. *Commun. ACM* 43, 9 (2000), 63–65. DOI: <http://dx.doi.org/10.1145/348941.348990>

[49] M. Six Silberman. 2016. Reading Elinor Ostrom In Silicon Valley: Exploring Institutional Diversity on the Internet. In *Proceedings of the 19th International Conference on Supporting Group Work (GROUP '16)*. Association for Computing Machinery, New York, NY, USA, 363–368. DOI: <http://dx.doi.org/10.1145/2957276.2957311> event-place: Sanibel Island, Florida, USA.

[50] Aaron Springer and Henriette Cramer. 2018. "Play PRBLMS": Identifying and Correcting Less Accessible Content in Voice Interfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA. DOI: <http://dx.doi.org/10.1145/3173574.3173870> event-place: Montreal QC, Canada.

[51] Selina Jeanne Sutton, Paul Foulkes, David Kirk, and Shaun Lawson. 2019. Voice As a Design Material: Sociophonetic Inspired Design Strategies in Human-Computer Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, 603:1–603:14. DOI: <http://dx.doi.org/10.1145/3290605.3300833> event-place: Glasgow, Scotland UK.

[52] Marie Louise Juul Søndergaard and Lone Koefoed Hansen. 2018. Intimate Futures: Staying with the Trouble of Digital Personal Assistants through Design Fiction. *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18* (2018), 869–880. DOI: <http://dx.doi.org/10.1145/3196709.3196766>

[53] Joshua Tanenbaum. 2014. Design Fictional Interactions: Why HCI Should Care about Stories. *Interactions* 21, 5 (Sept. 2014), 22–23. DOI: <http://dx.doi.org/10.1145/2648414>

[54] Mark West, Rebecca Kraut, and Han Ei Chew. 2019. *I'd blush if I could: closing gender divides in digital skills through education*. Technical Report. UNESCO, EQUALS Skills Coalition. <https://unesdoc.unesco.org/ark:/48223/pf0000367416.locale=en>

[55] Richmond Y. Wong, Nick Merrill, and John Chuang. 2018. When BCIs Have APIs: Design Fictions of Everyday Brain-Computer Interface Adoption. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, New York, NY, USA, 1359–1371. DOI: <http://dx.doi.org/10.1145/3196709.3196746> event-place: Hong Kong, China.

[56] Matthew Wood, Gavin Wood, and Madeline Balaam. 2017. “They’re Just Tixel Pits, Man”: Disputing the ‘Reality’ of Virtual Reality Pornography through the Story Completion Method. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5439–5451. DOI: <http://dx.doi.org/10.1145/3025453.3025762> event-place: Denver, Colorado, USA.

[57] Nicole Yankelovich, Gina-Anne Levow, and Matt Marx. 1995. Designing SpeechActs: Issues in Speech User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '95)*. ACM Press/Addison-Wesley Publishing Co., USA, 369–376. DOI: <http://dx.doi.org/10.1145/223904.223952> event-place: Denver, Colorado, USA.