Contextual Factors Affecting Information Sharing Patterns in Technology Mediated Communication

by

Anastasia Kuzminykh

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Computer Science

Waterloo, Ontario, Canada, 2020

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Dr. Carman Neustaedter Professor, School of Interactive Arts and Technology, Simon Fraser University
Supervisor:	Dr. Edward Lank Professor, David R. Cheriton School of Computer Science, University of Waterloo
Internal Member:	Dr. Daniel Vogel Associate Professor, David R. Cheriton School of Computer Science, University of Waterloo
Internal Member:	Dr. James Wallace Associate Professor, School of Public Health and Health Systems; cross appointed to David R. Cheriton School of Computer Science, University of Waterloo
Internal-External Member:	Dr. Stacey Scott Associate Professor, Department of Computer Science, University of Guelph; Adjunct Professor, Department of Systems Design Engineering, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This dissertation includes first-authored peer-reviewed material that has appeared in conference and journal proceedings published by the Association for Computing Machinery (ACM). The ACM's policy on reuse of published materials in a dissertation is as follows[1]:

> *Authors can include partial or complete papers of their own (and no fee is expected) in a dissertation as long as citations and DOI pointers to the Versions of Record in the ACM Digital Library are included.*

The following list serves as a declaration of the Versions of Record for works included in this dissertation:

Chapter 2
Anastasia Kuzminykh and Edward Lank. 2016. "People Searched by People: Context-Based Selectiveness in Online Search." In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. Association for Computing Machinery, New York, NY, USA, 749–760.
https://doi.org/10.1145/2901790.2901853

Chapter 4
Anastasia Kuzminykh and Sean Rintel. 2020. Classification of Functional Attention in Video Meetings. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–13.
https://doi.org/10.1145/3313831.3376546

Chapter 5
Anastasia Kuzminykh and Edward Lank. 2019. How Much is Too Much? Understanding the Information Needs of Parents of Young Children. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 3, 2, Article 52 (June 2019), 21 pages.
https://doi.org/10.1145/3328923

Chapter 6
Anastasia Kuzminykh, Jenny Sun, Nivetha Govindaraju, Jeff Avery, and Edward Lank. 2020. Genie in the Bottle: Anthropomorphized Perceptions of Conversational Agents. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–13.
https://doi.org/10.1145/3313831.3376665

[1] See https://authors.acm.org/author-services/author-rights accessed on June 5, 2020

Abstract

In this thesis, we investigate how and what contextual factors affect user's information sharing. We build our work on six individual research projects which cover a variety of systems (search engines, social network sites, teleconferencing systems, monitoring technology, and general purpose conversational agents) in a variety of communication scenarios with diverse relationships and dispositions of users. Alongside detailed findings for particular systems and communication scenarios from each individual project, we provide a consolidated analysis of these results across systems and scenarios, which allows us to identify patterns specific for different system types and aspects shared between systems.

In particular, we show that depending on the system's position between a user and an intended information receiving agent – whether communication happens through, around, or directly with the system – the system should have different patterns of operational adaptation to communication context. Specifically, when communication happens through the system, the system needs to gather communication context unavailable to the user and integrate it into information communication; when communication happens around the system, the system should adapt its operations to provide information in the most contextually suitable format; finally, when a user communicates with the system, the role of the system is to "match" this context in communication with the user.

We then argue that despite the differences between system types in patterns of required context-based adaptation, there are contextual factors affecting user's information sharing intent that should be acknowledged across systems. Grounded in our cumulative findings and analysis of related literature, we identify four such high-level contextual factors. We then present these four factors synthesized into an early design framework, which we call SART according to the included factors of space, addressee, reason and time. Each factor in SART is presented as a continuum defined through a descriptive dichotomy: perceived breadth of communication space (public to private); perceived specificity of an information addressee (defined to undefined); intended reason for information sharing (instrumental to objective); and perceived time of information relevance and life-span (immediate to indefinite).

Acknowledgements

I would like to thank all the people who made this thesis possible, including my thoughtful and patient mentors, my family, and of my dear friends who were always there to support me.

Dedication

I dedicate this thesis to my late grandparents, Dr. Vladimir V. Tyazhelov and Dr. Valentina G. Tyazhelova, whose approach to science and views on the role of knowledge in the society will always guide me in my research and my life.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Research presented in this thesis is motivated by the effort to support information sharing within communication processes, as well as further handling of this information, through its contextual interpretation. We suggest that, ideally, interpretations should be grounded in the recognition of the user's information sharing intent, which in turn is defined by the information sharing context. Consequently, this motivates a need for understanding the mechanisms of information sharing context effects on the user's intent. In this chapter, we first introduce a detailed description of the research space for this thesis and present our research questions (section 1.1). We then provide an outline of the thesis structure and summarize the thesis contributions (section 1.2), grounded in six research projects.

1.1 Research Space

The goal of communication is to exchange information. Furthermore how and what people communicate depends on context – their own situation and what they assume about the situation of a recipient [469].

Communication, as a process, consists of specific elements [425, 44], including *a sender* who produces *a message*, *a channel* used to share the message, and *a recipient* of the information, for whom the message is intended. The sender encodes the information in a message, sends the message through the channel, and the recipient decodes the received message. Respectively, the meaning of the shared information (i.e. the message) is symbolic [44] and encoded to be interpreted by a particular *recipient* according to the sender's sharing *intent*, where *intent* refers to the expected interpretation and use of the information by a receiving agent, whether this agent is a human or a computer system.

In today's world, communication frequently involves computers, either as the channel through which communication occurs, as a source of additional information in support of communication, or as a target for communication. Therefore, when I consider computation's role in communication, I differentiate systems depending on their "position" in communication. In particular, I view these system types as forming a continuum (Fig.1.1) from communicating *through technology* (e.g. videoconferencing systems, social media, etc.), to communicating *around technology* (e.g. a large digital display around which users collaborate, sensing devices for group activities, etc.), to communicating directly *with technology* (e.g. conversational agents, social robots, etc.).

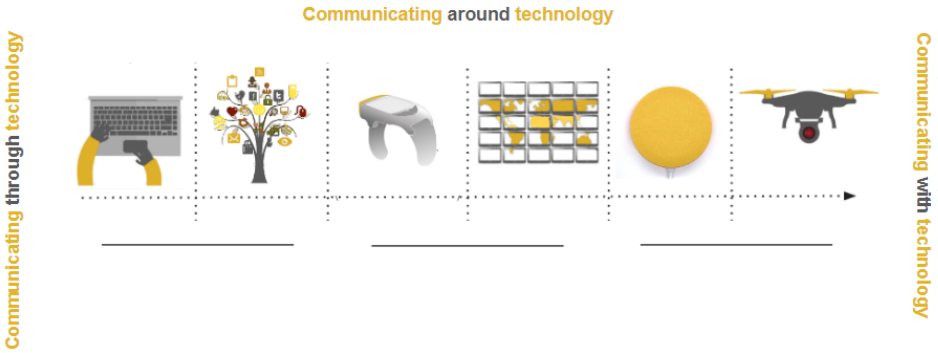

Figure 1.1: A continuum of system types with mapped examples from the author's work (online platforms [250, 253, 255]; social networks; wearable systems [251]; multi-user displays [494]; voice assistants [254]; social drones [249]).

Regardless of the role that computation plays in communication, the success of communication is predicated upon the receiver's recognition of a sender's intentions [24, 186, 419]. The sender's sharing *intent* (their expectations on the recipient's information interpretation and use) is defined by the context of the information sharing [206, 295]. For example, information might be considered private in some contexts, but public in other contexts [436, 296], meaning that its "secrecy" is a function of a particular information sharing context [355, 356].

Thus, the meaning of shared information is not a fixed set of properties within the shared data itself, but is, instead, grounded in the *co-existence* of *context-independent* information properties (activated on all occasions), and *context-dependent* properties (only activated by relevant contexts) [35], and data becomes information only once it is placed in context [157]. Then, an "objective" approach [206] to the interpretation of information as 'signals' [354] is not sufficient to guarantee intended interpretation [436, 296, 485], and

information should be interpreted "situationally" [206], with regard to the context of the communication [354].

Ideally, systems should be able to index, store, retrieve, and represent information in line with the sender's intent. However, intent is challenging for systems to infer because it is an internal cognitive process, i.e. is something the system does not have direct access to. One approach to inferring the user's intent is to view it as a specific case of the problem of abduction [453, 466], or reasoning to the best explanation. Given we have established that the sharing intent is formed according to the context of communication, it is possible that, by observing the sharing context, we can better infer the intent and related expectations on information interpretation and treatment. Although the intent itself is not directly accessible to the system, the system can access relevant context to try to infer the intent of the user when communicating.

To investigate the role that context can play in inferring intent requires, first, an understanding of how context affects a user's intent, and second, how this context can be understood by the system. This motivates the following **research questions**:

- (RQ1) How does the context of information sharing affect users' intent (willingness to share information and expectations on information interpretation, use and treatment)?
- (RQ2) What specific factors create this information sharing context?

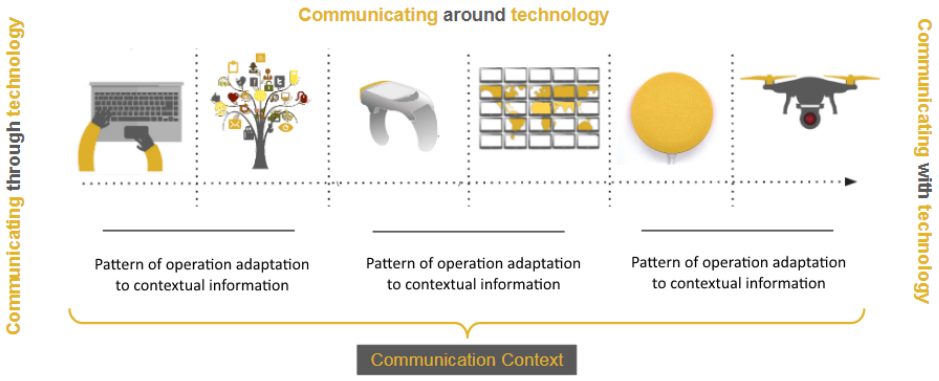

Figure 1.2: A continuum of system types. Pattern of operation adaptation differ for each system type (RQ3), while relevant contextual factors are shared across systems (RQ4).

The diversity of systems supporting communication means that these systems have different target functionalities (e.g. providing social content, search of information, task

assistance, etc.), and so they would incorporate context in different ways to support information sharing [29, 464]. However, we hypothesized that systems, grouped according to their "position" in communication (Fig. 1.2), will share common patterns of adaptation to the context. Additionally, we hypothesized that there are aspects of context that should be considered across systems. This motivates two additional **research questions**:

- (RQ3) What are the patterns of required adaptation to information sharing context for different types of systems?
- (RQ4) What high-level contextual factors that define user's information sharing intent should be considered across the different system types?

In essence, we expect, first, similarities within each system type in what systems should be doing based on context (how they should be adapting), and second, while specific low-level contextual data considered as relevant to the particular interaction might vary, the high-level contextual factors should be shared between the system types (Fig.1.2).

To summarize, in this section, we suggest that successful interpretation of shared information requires recognition of the sender's sharing *intent* (their expectations on the understanding and use of information) which is defined by the *context* of the communication. Although a system does not have direct access to the user's intent, it may be able to infer it based on the accessible contextual factors available. Therefore, in this thesis, I explore the factors that create information sharing context (RQ2) and their effects on the sharing intent (RQ1). Due to the diversity of communication systems, we also introduce a continuum to differentiate systems types according to their position in communication, i.e. as supporting communication through, around, or with the system. We hypothesize different patterns of adaptation of system operation based on communication context (RQ3) with the overall possibility of identifying a set of contextual factors relevant across the systems (RQ4).

1.2 Contributions

1.2.1 Research Strategy

In this thesis, we present six targeted research projects. The set of individually covered systems is motivated by the suggested approach to system types, which differentiates between communication through, around, and directly with technology. The choice of individually covered communication scenarios was guided by an effort to investigate diverse

sender-receiver relationships and, correspondingly, diverse information sharing contexts. Specifically, included in this thesis, we explore:

1. personal online content management, online persona, and online reputation management through search engine systems – Chapter 2 [250];
2. interpretation of the information which includes children, shared by parents through social media with their social network – Chapter 3;
3. impact of user's interpretive communicative attention actions on engagement in hybrid work meetings (in which participants are distributed unevenly over teleconferencing endpoints) – Chapter 4 [252];
4. parents' use of monitoring technologies for children, including information needs, motivations and privacy concerns, the use-cases for this information, and desired information format – Chapter 5 [251];
5. structure, interconnectedness, and agent-specific differences of the anthropomorphized behavioural and visual conceptions of conversational agents – Chapter 6 [254];
6. user's conceptualization of conversational agent work process, including envisionment of possible tasks, related performance expectations, and associated concerns regarding the "expectedly unexpected" resource dynamics – Chapter 7.

Projects 1 and 2 represent communication *through* the system, projects 3 and 4 communication *around* the system, and projects 5 and 6 *with* the system. This organization of research projects allows for subsequent co-joint analysis of the patterns of required adaptation to information sharing context for different types of technology (RQ3). Furthermore, co-joint analysis of particular contextual factors and their effects discovered within each of the individual projects (RQ1 and RQ2), allows us to draw a set of contextual factors that should be considered across system types (RQ4).

1.2.2 Structure and Contributions

We organize the presentation of our contributions in four parts (Fig.1.3). First, we discuss synthesized patterns of required adaptation to information sharing context, each grounded into two individual research projects:

- when communication occurs **through the system** (Part I, chapters 2 and 3),
- when communication occurs **around the system** (Part II, chapters 4 and 5), and
- when communication occurs **with the system** (Part III, chapters 6 and 7).

Due to the diversity of individually covered systems and communication scenarios, targeted review of related literature in included in each chapter.

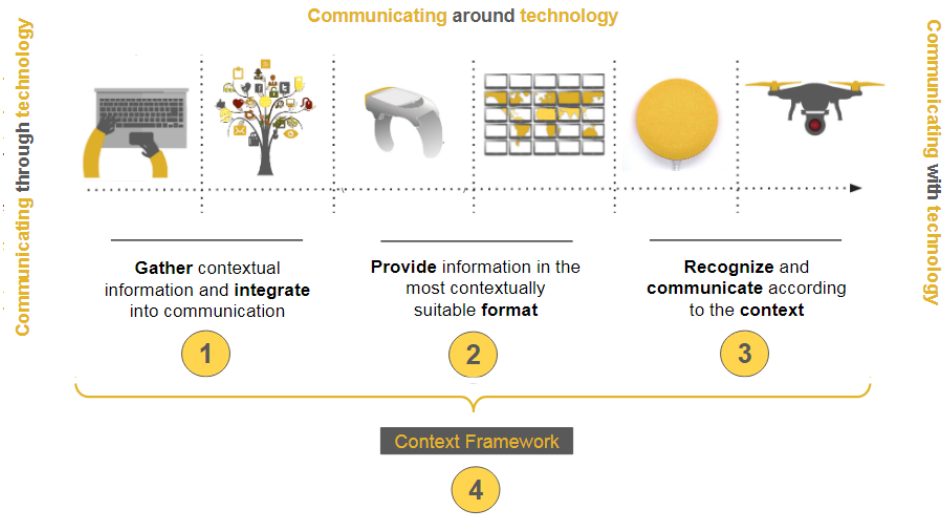

Figure 1.3: The structure of this thesis, organized according to four parts.

We then present the general discussion of literature on context in information sharing, as well as the synthesized design framework (Part IV, chapter 8), which encompasses four contextual factors that should be considered across system types. We define each factor through a descriptive dichotomy, and discuss their effects on user's information sharing intent, grounded in our findings and related literature.

Communication through the system

In Part I of this thesis, we demonstrate that when communication occurs **through the system**, the role of the system becomes *to gather communication context unavailable to a communication endpoint* and *to integrate it into communication process.*

Research project 1 (chapter 2) explores personal online content management with regard to search engine systems. We examine the search practices and perceptions of online search on one's name, including both how individuals manage online content about themselves and how they interpret and contextualize information about others.

Our results, first, demonstrate a desire for agency over what information is provided on one's name, both from searchers and searched individuals, grounded in a need for contextualization of information. In particular, searched people are frustrated with the notion that

the prominence of certain contents and whether or not this information is found in a given searching context is outside of their power. Complementary to these results, searchers express a desire to constrain their exploration of online content to a subset of information deemed relevant to a particular context when searching others. Thus, we demonstrate that search context heavily influences the type of content expected from a system to provide. This context specifically includes social factor – the nature of relationship between searchers and searched individuals, and temporal factor – the age of the information. Interestingly, while age of the information consistently affects its perceived relevance within each given social context, e.g. professional, social, and romantic, the specifics of temporal relevance of information (the clusters of age of interest) vary between different social context.

To summarize, our synthesised results from **project 1** show that there is a prevalent desire for agency, both from searchers and from searched individuals, to be able to prioritize and categorize content to constrain the accessible information to the specific goal of the search. More specifically, users are aiming to consider social and temporal context in assessment of relevance of online material, however, find little support for such contextualization from the search system. To illustrate how this desire for contextualization of search results can be addressed by information retrieval system, we provide a design sketch of a "personal search page", where search results on an individual are provided.

In **research project 2** (chapter 3), we focus on social media systems and investigate a distinct social media behavior labeled sharenting – the phenomenon of sharing information about one's child online. Such parent's practice of online information sharing presents a number of issues associated with the complex balance between parent and child identity representations in online posts.

We explore factors that are guiding one's perception of a sharenting post as being about a parent or about a child. We refer to this decision as the designation of a protagonist in a sharenting post. We examine the designation of a protagonist from two perspectives – of a sharing parent (an author), and of a viewer of a post.

We first show the contextual ambiguity and dynamic nature of a protagonist assignment, i.e., a perceived protagonist in the same post changes over time and between contexts. Specifically, considerations regarding child's safety and their future privacy revolve around a child's identity. However, when a post is considered within the communicative frame of reference, within which it was intended (sharing with parent's network), the identity of the protagonist shifts toward the parent. We then identify four overarching factors affecting the designation of a protagonist, both for posters and for viewers of sharenting information, including locus of sharing motivation, agency in sharing, social connection to the information recipient, and agency of the captured action.

Overall, **project 2** illustrates that sharenting is a prime example of posting behaviour where content perception – particularly, protagonist perception – evolves with time and within different contexts. However, these perception of the protagonist in each case can be traced to a set of particular affecting contextual factors, accessible for analysis.

Consolidated analysis of **project 1**, focusing on search engines, and **project 2**, focusing on social media, demonstrates how the systems, through which communication occurs, can support this communication by augmenting it through gathering inaccessible contextual information relevant for information interpretation. For instance, **project 1** highlight the user's need to prioritize content to interpret it according to the social and temporal aspects of context relevant to the given search. Since this social and temporal context of information perception is unavailable to the poster at the moment of information sharing, the search engine is expected to adapt information retrieval according to the state of corresponding contextual factors at the relevant moment. In a similar vein, **project 2** shows how specific social and temporal contexts of consideration of sharenting posts affect the changes in the perceptions of who this post is about (the protagonist in the post). Such context-dependant ambiguity of a protagonist suggests the need to appropriately segregate parent's and child's identity in information representation. For example, one potential solution to this problem is for social networks and information retrieval systems to monitor how information from sharenting posts is re-used, and to retrieve it in association with appropriate identity's name – adapt information retrieval and presentation according to the relationship between parent's original sharing intent and context of the retrieval.

Communication around the system

In Part II, we suggest that when communication occurs **around the system**, the role of systems, in this case, becomes *to provide information in the most contextually suitable format*. We demonstrate this pattern of required adaptation through projects 3 and 4.

In **research project 3** (chapter 4), we explore video-mediated participation in hybrid work meetings, in which participants are distributed unevenly over teleconferencing endpoints. Specifically, we research factors affecting participants' engagement by analysing prioritization and sense-making of visual attention during video meetings at work.

Synthesizing our findings, we first identify three *categories of attention* that a video meeting feature might support: attention as *Direction* – purely observable attention that constitutes a process, but not a recognized purpose of it (e.g. Alice is looking at Bob), attention as *Action*, which reflects attention processes with a specifically recognized purpose (e.g. Alice is addressing Bob), and attention as *State* – the overall sense of engagement,

which reflects only purpose, but no particular observable process (e.g. Alice is attentive). The purpose of attention, in this case, refers to its contextually interpreted function. We demonstrate that for attention as *State* such contextually interpreted function is a general sense of engagement ('be mentally there'), and attention as *Action* supports three contextually interpreted functions: – gathering of information ('look to see'), communicative signaling ('look to be seen'), and following dynamic processes ('look to be a part of a process'). Importantly, our results reveal that the accessibility of means for the efficient behavioural engagement (performance of attention as *Action*) significantly contributes to the participant's overall sense of engagement, and these effects are precipitated by the limitations that remote participation puts on the three functions of visual attention. Thus, to be able to keep remote participants engaged, the system should be able to support attention as *Action*, which requires system's operation adaptation based on augmenting pure visual signal input of attention as *Direction* with an interpretive meaning according to one of three functions of attention as *Action*.

While traditionally teleconferencing systems would be considered as supporting communication *through* the system, we argue that, in case of hybrid meeting scenarios, they represent communication *around* the system. In particular, as we discussed, our results highlight the effects of the restrictions, that a teleconferencing system imposes on the user's ability to fully communicate contextual meaning of their actions. In essence, these restriction are based on user's inability to position their action corresponding to spatially organized objects or people. For example, addressing a person with one's gaze implies positioning the gaze direction in correspondence to spatial location of this person (signaling function of Attention as Action). Similarly, following a particular thread of conversation when multiple branches of discussion appear in a room requires user's spatially positioning themselves according to the thread of interest (function of following dynamic processes). Thus, we find that actions which in non-mediated scenarios gain their contextual meanings from spacial disposition according to the other agents in the environment, loose their contextual meanings due to the limitations of traditional video-mediation.

Research project 4 (chapter 5) demonstrates that tracking and monitoring technology for children, in fact, also present an example of communication *around* the system. In this project, we investigate parent's information needs, motivations, privacy concerns, and use-cases for monitoring information to develop design approaches in light of parent's resistance to pervasive monitoring.

Based on accumulated findings from three studies, this project, first, shows the structure of monitoring information that parents seek, organized into four categories: *routine* information, *health* information, *school* information, and *social-emotional* information. We then analyze parents' underlying motivations for seeking the information and demonstrated

how these motivations are affected by parents' beliefs about privacy, trust-based relationships with their child, and a need for work-life balance. We show that the information is predominantly desired when an action is required from a parent, and, consequently, identify the common use-cases for this information: *adjusting routine activities*, *performing emergency activities*, and *building communication and personal relationship with a child*.

Our results suggest that consideration of parents information needs in context of the intended information use yields a promising approach in addressing the tension between the need to monitor to identify required actions and resistance to pervasive monitoring.

Both **project 3** and **Project 4** suggest that when communication occurs **around the system**, this system should adapt to the information sharing context by providing the information in the contextually suitable format. **Project 3** shows that a teleconferencing system should adapt its operation based on augmenting pure visual signal with an interpretive meaning of attention action (their function). For example, system features aiming to support information gathering function should involve timely augmentation of information streams, e.g. narrating, zooming-in, or focusing system's view or bandwidth. Supporting communicative signaling function, on the other hand, requires some form of notification about one's actions delivered to other meeting participants. **Project 4** shows the role of processing and presenting the information shared with a monitoring system according to the context of its intended use. For instance, on a typical day, parents don't want a constantly available stream of information, but instead they prefer summary information of what has happened during the day apart. Parents then would use this information to adjust their behavior when with a child later that day, or as a basis for shared activities.

Communication with the system

Finally, in Part III, we show that when communication occurs **with the system**, the role of this system is to recognize user's context and *to "match" this context in communication with the user*. We ground our argument into results from projects 5 and 6.

Research project 5 (chapter 6) explores anthropomorphized behavioural perceptions and visual conceptions of conversational agents and identifies patterns of such perceptions and their association with specific designed features of the systems.

Drawing from a comparative analysis of three conversational agents – Alexa, Google Assistant, and Siri, – we first demonstrate that the perceptions of an agent's character (anthropomorphized behavioural perceptions) are structured according to five categories: social approachability of an agent, their sentiment toward a user, their professionalism, intelligence, and individuality. We also discuss the differences between agents' perceived

characters. For example, Google's character tends to be described as approachable, closely followed by Alexa, but Siri is more commonly described as distant. Similarly, Google and Alexa are often seen as more genuine and caring, while Siri's sentiment is seen as disingenuous and inconsiderate. We then explore visual conceptions of the agents – how users think an agent would look like if they were human. We analyze particular appearance elements (e.g. body build and age, facial expressions, hair, and clothing style), describe the structure of each of these categories, and discuss the consistent specifics assigned to each agent. For instance, Alexa tends to be seen of average height or shorter, older than other agents, with darker, wavy, and worn down hair, and casual clothes highlighting either a tendency to be unnoticeable or cozy, home-associated nature. Google is seen of average height or taller, with lighter hair often worn up, (associated with higher professionalism), and their clothes reflect perception of them as a busy, stressed, or rather "awkward" person. Siri is commonly described as being of an average height, younger than the other agents, and wearing either casual but fashionable clothes highlighting a daring personality with accentuated femininity. Finally, we demonstrate that the character perception is self-reportedly motivated only by the designed features, while visualizations are motivated by some design features, but are predominantly associated with the character's perceptions.

The combined results from this project demonstrate that the anthropomorphized behavioural and visual perceptions of agents yield structural consistency and are linked with each other and system features. Furthermore, the richness of socially meaningful interpretations, including users' extreme tendency for displaying biases and social stereotypes, highlights the importance of considerations of user's perceptions during their communication and information sharing with the system.

Research project 6 (chapter 7) also focuses on general-purpose conversational agents, but explores user envisionment of the possible tasks, related performance expectations, and associated concerns regarding the "expectedly unexpected" resource dynamics.

The findings from this project are synthesized as a conceptual model of the conversational agent work process. First, the model includes a classification of interaction types for conversational agents' performance, based on a combination of users' involvement in the primary activity and the nature of interacting entities (human, system, environment). For instance, two "parent" interaction types are identified according to the level of the user's involvement in the primary activity: "assistant-user" interactions ("help me to do" tasks that facilitate the user's performance of main activity) and "assistant-world" interactions ("do instead of me" tasks where main activity is performed by an agent). We then present a corresponding typology of envisioned conversational agents tasks, organized according to the interaction types they represent. The task categories were formed based on the interaction of four parameters: content, required device infrastructure, required system

intelligence, and the temporal aspect of the task. We continue by exploring the dynamic aspects of the task processes and reflect it through a schematic representation of two task execution flows associated with "assistant-user" and "assistant-world" interaction types from the perspective of resource dynamics and task outcomes. Finally, the fourth component of the model is a classification of users' concerns mapped onto the task execution flows based on the nature of concerns and their correspondence to system architecture and design components. The demonstrated associations between the problem spaces and their "place" in the task execution flow is the first systematic representation of user concerns in a dynamic context of a task performance.

Overall, the first two components of the model show functional differences between tasks, and dynamic aspects of the third and fourth components demonstrate associations between task functions, resource flows, and user's concerns. Such systematic understanding of the conversational agents' design space allows researchers to develop complex approaches to supporting user's mental model of these systems and to address adoption obstacles and guide privacy decisions through appropriate feedback and visualization mechanisms.

To summarize, both projects suggest that when communication occurs directly **with the system**, the role of this system is to recognize user's information sharing context and communicate with the user according to this context. For instance, the depth and richness of social interpretations of conversational agents, demonstrated in **project 5**, highlights the necessity to consider social context for the interaction, as implied by the user, e.g. perception of an agent as "cozy and homey" could provoke different information sharing patterns compared to a "distant, cold, strictly professional" perception. **Project 6**, in turn, shows the potential for a system to appropriately match its feedback to the user's contextual expectations of recourse dynamics and task performance.

Contextual Factors as Design Framework

The consolidated analysis of our findings on *contextual factors* of information sharing (RQ2) and their effects on user's *information sharing intent* (RQ1) when communicating *through*, *around*, and *with* the system (RQ3) allowed us to draw a set of *contextual factors* that should be considered across system types (RQ4). We synthesize them as a design framework (Part IV, chapter 8), which we refer to as **SART** based on the first letters of each included contextual factor - **space**, **addressee**, **reason** and **time** where each context factor (heuristic) is defined through a descriptive dichotomy.

- The **space factor** is represented by the public space to private space continuum and refers to the perceived breadth of the audience.

In individual research projects, we have observed the effects of space interpretation, for example, in **project 1**, (chapter 2). which revealed that both searchers and searched individuals want to receive prioritized and categorized content to *constrain* the accessible information to the specific context of the given search – targeted narrowing of the breadth of the audience. Another example of effects of space interpretation was observed in relation to sharenting behaviours (**project 2**, chapter 3), where parent's considerations of their audience as more "private" and narrow (the audience *intended* by a sharing parent) would provoke looser sharing practices, while perceptions of potential *broader audience* provoke stricter sharing rules. Thus, higher readiness to share the information is associated with parent's perceptions of a "degree of controll" over their audience. Furthermore, these mechanisms can be observed through the analysis of parents' underlying motivations for seeking information from monitoring technologies (**project 4**, chapter 5). While such information about a child is seen as highly sensitive and motivates increased attention to security and privacy consideration, its intended use within *private* (even rather *intimate*) relationship between parent and child increases parents' willingness to share information through technological mediation. The opposite effect of space interpretation appears in privacy considerations in interactions with conversational agents (**project 6**, chapter 7), becoming particularly relevant when an agent performs direct communication with *others* (e.g., calling to make an appointment), which invokes specific expectations for an agent to *differentiate* between "private" and "public" information.

- The **addressee factor** is represented by the continuum from "defined" to "undefined", referring to the degree of specificity of the expected information recipient.

Although **space** and **addressee** factors are conceptually related, we argue that they represent two distinct factors. Specifically, as we mentioned, the **space factor** refers to the perceived breadth of the audience for the shared information. However, even when the audience is perceived to be broad (as opposed to targeted), the specifics of this audience might be rather defined. For example, in relation to personal information retrieved by search engines (**project 1**, chapter 2), we identified three distinct social contexts – one's professional, social, and personal online personas. While inability to highlight content related to one of these personas and silent it for other contexts causes the expected breadth of the audience, this audience is still rather defined, leading to the user's desire to ensure the existence and accessibility of certain content (e.g. professionally), regardless of the audience breadth. On the other hand, expected narrow audience might be "undefined". Consider communication with unfamiliar person on the phone – the audience is narrow and targeted, yet, has lower degree of specificity, i.e. the user has limited contextual information to form appropriate expectations on interpretation and use of the shared information.

Another argument for disentangling **space** and **addressee** factors is that there are different design mechanisms that can be applied towards each of them. For instance, when communication occurs directly *with* the system, the degree of specificity of **addressee** – how defined it is – can be approached by provoking anthropomorphization. Indeed, in **project 5** (chapter 6), we find that anthropomorphized behavioural and visual perceptions of agents are structurally consistent and have rich socially meaningful interpretations, e.g. social stereotypes, thus, increasing the perceived **addressee** specificity.

Project 3 (chapter 4) on video-mediated participation in hybrid work meetings also demonstrates how lower degrees of contextual "definitiveness" of information recipient affect one's engagement in a communication process. For example, lacking ability to form contextual interpretations of remote participants' states and actions (restrictions to their signaling function) lead to their social exclusion from an active meeting process, and their lacking ability to collect contextual information to form appropriate "definitiveness" of information recipient (restricted function of social information gathering) causes their lower motivation to engage. Especially prominently the effects of specificity of an addressee can be seen in relation to sharenting behaviours (**project 2**, chapter 3) – within the intended communication context (with a parent's social circle), a protagonist tends to be assigned to a parent, amplifying parent's desire to share the information. However, when the imagined audience becomes less defined (the present and future "unintended" audience), the expected shared communication context is lost, and the interpretation of information changes both for the sharing parent and for the viewers of this information. Furthermore, the defined perceptions of "unintended" audience also affect sharing patterns, e.g. limited sharing due to the anticipated malicious intentions.

- The **sharing reason factor** refers to the primacy of the intended information perception, stretching from "instrumental" sharing to receive some service benefit to "objective sharing" as the main goal of the activity.

Arguably, any intended information sharing can be considered instrumental since it is goal-oriented. Here, however, we refer to information sharing as "instrumental" if its performance is secondary within the main task. Users' willingness for "instrumental" information sharing is defined by the anticipated boundaries of the main task. "Objective" information sharing, on the other hand, implies that its performance is primary to the task (e.g. a post on social network), and willingness for "objective" sharing is defined by the expectation of its correct interpretation and corresponding social benefits.

Parents' use of monitoring technologies (**project 4**, chapter 5) is an example of "instrumental" sharing. The child's information collected by the system (shared with the system) is perceived as highly sensitive, but parents are willing to provide this information if it

properly addresses its use-cases. Correspondingly, the retrieval of this information is only desired in association with the sharing intent (when an action is required) and in suitable format – summarized data on a regular day and detailed data in case of emergency. Another example of "instrumental" information sharing appears in perceived resource flows in interactions with conversational agents (**project 6**, chapter 7) – concerns regarding the resource flow seem to be lower for user-initiated tasks, and particularly for "assistant-user" interactions, when conversational agent facilitates the performance of the primary activity, while the user actually performs it ("help me to do" tasks). On the contrary, users express significantly more concerned beliefs regarding an agent's unauthorized and uncontrolled access and unauthorized passing of their resources – information that was shared with the system for the purpose of a particular task but feared to be used outside of its scope.

"Objective" information sharing is reflected, for example, in sharenting practices (**project 2**, chapter 3) and users' practices of online reputation management (**project 1**, chapter 2). These projects demonstrate how the expectations of interpretations and social relevance of shared information affect user's willingness to share or to limit information dissemination. Similarly, the analysis of technology mediated participation in work meetings (**project 3**, chapter 4) further illustrates the facilitating role of expected social interpretations in willingness for objective information sharing. For instance, consider the limitations imposed by video-mediation on the signaling function of attention actions and the effects of these limitations on communication engagement.

Finally, anthropomorphized perceptions of technology, e.g. conversational agents (**project 5**, chapter 6), reveal an interesting question on the balance and relationship between "instrumental" and "objective" information sharing within one system. In particular, the social nature of the environment which these systems occupy in combination with the discovered richness of socially meaningful interpretations of their interactions provokes users for "objective" sharing. Indeed, conversational agents are capable of supporting diverse purely social, non-task-based interactions (e.g. [378]). On the other hand, main functionality of conversational agents – personal assistance – implies "instrumental" sharing. Thus, it is system's responsibility to correctly recognize the acceptability and appropriateness of use of information based on the contextual factor of sharing reason.

- The **time factor** refers to the anticipated length of the information life-span, stretching from perceived "immediate" to perceived "indefinite" information life-cycle.

Overall, the forthcoming to the "immediate" end of the spectrum tends to provoke easier information sharing, while expectations of its longer life-time motivate users to limit their sharing activities. In part, it is related to the increasing chances of the information to be misinterpreted (connection to the **addressee** factor), to loose its descriptive relevance

over time, as well as to be used for unintended purposes (connection to the **reason** factor).

For example, the contextual effects of the time factor are illustrated by one's sharing practices with regard to their online reputation management (**project 1**, chapter 2), where the age of the information plays a significant role in the assessment of its relevance for a particular social context, and users express a desire to be able to prioritize content accordingly. Similarly, the analysis of protagonist assignment in sharenting posts (**project 2**, chapter 3) shows that the "future" perceptions and interpretations of the information are expected to shift towards the child's identity, demonstrating how these considerations affect parent's willingness to share the information.

The associations between the **time factor** and the intended "instrumental" **reason** for the information sharing is revealed, for example, through the analysis of privacy concerns in interactions with conversational agents (**project 6**, chapter 7), e.g. perceptions of agents "storing" the information. Interestingly, however, it did not appear in our data on parents' information needs and use monitoring technologies (**project 4**, chapter 5), potentially due to the perceived "immediacy" of the information use and relevance (practical focus of use-cases for routine activities).

To summarize, based on the accumulative analysis of findings from six research project, the **SART** framework brings together four contextual factors – **space**, **addressee**, **reason** and **time** – each defined as a continuum through a descriptive dichotomy. These four contextual factors are interdependent, however, drawing distinctions between them is an important step toward systematic and functional understanding of context that affects information sharing intent. In Part IV of this thesis we present an extended discussion of each of the four contextual factors, grounded both in our findings and in related literature.

Discussion and Implications

Finally, in chapter 9, we conclude by discussing directions for future work, and the role of this research in advancing continuous efforts within the paradigms of human-computer symbiosis [273], human-computer collaboration [466], and proactive computing [465].

Part I

COMMUNICATION THROUGH TECHNOLOGY

Figure 1.4: Pattern of operations adaptation in communication through the system. Examples from our work (left [250]), right [255]).

In the first part of this thesis, we analyse communication *through* the systems. Corresponding type of systems represents technology and platforms supporting communicating with other people, e.g. social networks platforms, audio and video conferencing systems, or, for example, file sharing services. The mediated nature of user's communication through these systems, whether synchronous or asynchronous, commonly implies that information sharing becomes distributed in space and/or time, which prevents users from acquiring the necessary context to form appropriate expectations and interpretation of the shared information [203, 277, 529, 511, 71, 280]. Who is reading user's posts? Who can see and hear the user when they are on the screen? How are people reacting to the shared information? When and in what situation do others receive the information? One users' strategy to resolve these uncertainties is to create a conceptualization of people with whom they communicate [4], so called imagined audience [278], as well as a conceptualization of the expected situations for the information consumption. However, the actual audience and their situation of receiving this shared information would often differ from the imagined one [280, 320, 93, 95, 302], loosening interpretive significance of the information sharing context (e.g. [481, 301, 225, 115]). Another strategy is to limit disclosure and communication participation [9, 448, 261, 344, 432, 253], i.e. to avoid information sharing, which by definition defeats the purpose of communication.

We suggest that when communication occurs **through the system**, the system's general operation adaptation pattern should be *to gather communication context unavailable to the communication endpoint* and *to integrate it into information communication* (Fig. 1.4). We ground this suggestion into consolidated findings from two research projects, one focusing on the use of search engines in personal online content discovery and management (**project 1**, chapter 2), and one focusing on ambiguity of interpretation of content posted

on social media (**project 2**, chapter 3).

The focus of **research project 1** (chapter 2) is motivated by the observation that online content plays an important role in reputation management, however, once information is online, it quickly becomes part of the internet archive with little possibility for user's control over it. While one way to manage an online persona is to restrict the information provided in the online domain, there are significant benefits to having an online presence for one's social and professional life, and users tend to perceive it necessary to exchange personal information to realize their social goals and gain the social capital [140, 458, 336, **?**]. Furthermore, the information about an individual might be posted by third parties. In this project, we explore the search practices and perceptions of online search on one's name, including both how individuals manage online content about themselves and how they interpret and contextualize information about others.

We draw our results based on two consequent studies: an interview study that explores online practices including how and why participants share information and the types of searches participants perform seeking information on others, as well as an online survey that expands upon the data from our interviews on users' strategies for vetting, managing, and interpreting content online. In synthesizing our results, we demonstrate a desire for agency both from searchers and searched individuals, grounded in a need for contextualization of information according to social and temporal context. Specifically, users are aiming to constrain the accessible information to the specific goal of the search by considering social and temporal context in assessment of the relevance of online material, however, find little support for such contextualization of information from the search system.

Research project 2 (chapter 3)) is motivated by the challenges of identity representation in social media that include information on different individuals. We specifically focus on interpretation of a primary identity in information shared by a parent about their child online, a distinct social media behavior labeled sharenting [55, 64, 444, 98], since sharenting has been demonstrated to present considerable issues associated with the complex balance between parent and child identity representations [167, 64, 444, 114, 98]. In this project, we explore factors that are used to decide whether a post is about a parent or about a child, and we refer to this decision as the designation of a protagonist.

Through a set of semi-structured interviews with parents of young children, we explore self-reported factors taken into a consideration when assigning a protagonist to a sharenting post from two perspectives – as a sharing parent (an author), and as a viewer of a post. Our findings reveal the contextual ambiguity of a protagonist assignment, demonstrating how the identity of the protagonist of sharenting posts changes both over time and depending on perspectives taken when viewing the content. To gain insight into what

defines this ambiguity and dynamics, we then explore protagonist discrepancies between different contexts, and, consequently, identify four factors that influence the perception of the protagonist, which can potentially be leveraged to support identity management through contextual cues. These factors include: locus of sharing motivation – whose interests and social needs are intended to be addressed by the post, agency in sharing – an actual decision to share specific information and a necessity to acquire relevant permission, social connection to the information recipient – the social relationship between a viewer, the sharing parent, and the child, and agency of the captured action.

Both **project 1** and **project 2**, illustrate the issues rising from users' lacking ability to acquire the necessary context to form appropriate expectations and interpretation of the shared information, suggesting that systems, through which communication occurs, should support and augment this communication by gathering and integrating contextual information into user's communication. For instance, **project 1** highlight the need to prioritize content most relevant to the context of search according to the users' desire to interpret content with respect to social context and time. While this social and temporal context is unavailable to the poster at the moment of information share, the search engine can collect required contextual information at the moment of information retrieval and adapt information presentation accordingly. **Project 2** shows that depending on whether sharenting posts are considered from the perspective of current or future risk or from the perspective of communicative goals, perceptions of who a post is about (the protagonist in the post) changes. A sharenting social media post, as a communicative message, is intended – encoded – by the sharing parent for consumption – decoding – by some recipient. Then, the specifics of the audience intended by a sharing parent for the post determine parent's expectations on how, and, in particular, about whom this information is going to be perceived – within the intended communication context with a parent's social circle a protagonist tends to be assigned to a parent, but when the audience becomes less defined (the present and future "unintended" audiences), the expected shared communication context is lost, and a protagonist is more likely to be assigned to a child. Such context-dependant ambiguity of a protagonist suggests that the system should support appropriate identity segregation in information representation according to the context of information consumption.

The following two chapters present a detailed discussion of **project 1** and **project 2**.

Chapter 2

Project 1: Context-Based Selectiveness in Online Search

In this chapter, we look at a first instance of communication through technology. In particular, the focus of this chapter is an asynchronous communication between individuals using named search in a web search interface. We explore the desire for an ability for control of what content is displayed, both from the side of the searched individual and from the side of the searcher.

2.1 Introduction

> "...if I look at your Facebook profile and realize that you have kids! Or you're a woman even trying to get pregnant. That is not a reason for me not to hire you. I'm legally not allowed not to hire you for those reasons. So I'm putting myself in a particularly dangerous position even by finding that out... It is dangerous to even look at this stuff."
>
> *[P21], 12 years of hiring experience in software engineering industry*

As a society, we have begun to take note of the challenges associated with the preservation of privacy and the protection of reputation in light of widely available online content. The recent European Union Court of Justice decision to enforce on search engines the right of individuals to be forgotten is one way that, collectively, society has tried to enforce some limits on the persistence of online personal information [14, 450]. Alongside

legal remedies to allow individuals to control online content about themselves, researchers have also begun to struggle with the need to develop mechanisms that allow individuals to manage their online personas, i.e. to control the access to online information [25, 37, 265, 105, 124, 234, 306, 340, 426]. This research spans many aspects, including the usability of privacy settings on social networking sites [124, 4, 187, 243], theories of user behavior while online [234, 426, 314], and analyses of techniques people employ to maintain and correct online personas [306, 300, 104, 384]. Finally, researchers have explored the link between online content and an individual's reputation [306, 511]. Our interest is specifically within this domain of reputation management and techniques or interactions that can help individuals' manage online content with respect to its impact on reputation.

Reputation can be conceptualized as the common opinion that people have about someone or something, or the way in which people think of someone or something. Inherent in this definition of reputation is the notion of judgement. One's reputation is a construct informed by both direct and indirect information obtained from a variety of sources, and interpreted and contextualized by the perceiver to formulate an opinion. The online world, in both how accessible it makes data and in the persistence of data, has a huge potential to influence one's reputation [306, 300, 511]. While it is perceived to be necessary to manage one's reputation online [450, 306, 300], researchers also report that users find the process of online reputation management distasteful and disempowering, particularly when negative content impacts one's reputation [511, 104, 384, 503].

One way to manage an online persona is to carefully restrict the information provided in the online domain, i.e. one can simply not post, or restrict access to information [426, 103]. However, information about an individual is not always posted by that individual [258, 259]. News organizations, bloggers, acquaintances: The potential sources of content on an individual are myriad and suppressing all of this content is difficult prior to release. Suppressing or hiding content post release (as in the suppression of search results by search engines [14]) is one way that individuals can attempt to control information online, although this often fails [503]. Unfortunately, given the existence of tools such as the internet archive (archive.org/web), content can always be found. Assuming content exists and persists, are any other remedies available?

The concepts explored in our research are heavily influenced by the concept of selectiveness as articulated by Dourish and Anderson [131]. As we began to explore online content, online persona management, and online reputation, one characteristic we observed of much of the past research was that it viewed the individual about whom an opinion is being formed, i.e. the target of a search or the searched, as a source of data for the search engine [124, 4, 103, 106] rather than as an influencer – through the search engine - on the perception of the searcher [37, 306, 104]. But... one's reputation is a construct of others,

"the way in which people think of someone". This, then, lead us to ask a simple question: What about the searcher? Can we influence the way someone thinks about another person even in the absence of an ability to suppress information? Perhaps, stated another way, is there a way that we can tighten the communication channel between searched and searcher such that both become participants in the discourse surrounding online content, i.e. they both become users of the search engine serving information to a searcher?

Understanding the concept of selectiveness is, we feel, essential to understanding online content management, online reputation, and online privacy because we concur with Dourish and Anderson [131] that talking about secrecy, privacy, or the blocking of access to information in the generic sense is inappropriate. The secrecy or the limitation on information provision is both a function of sender and receiver, i.e. there is implicit selectiveness at work in every online interaction. What you share with one, in the moment, you may wish not to share with another. Therefore, we began by exploring how selectiveness was realized when individuals were disseminating content online.

In this project, we present our results exploring the search practices and perceptions of online search, including both how individuals manage online content about themselves and how individuals interpret and contextualize information about others. We describe two studies. First, we describe and present the results of an interview study that explores online practices including how and why participants share information and the types of searches participants perform seeking information on others. Next, we describe the results of an online survey that expands upon the data from our survey.

From our interview data, we find that, when individuals release data, they frequently think about the recipient of that information as a factor in contextualizing the interpretation of any online data. As well, we find that there exists a temporal saliency to online information and an appreciation of the fact that this temporal saliency is poorly supported by current search engines. Finally, when individuals search for data online, then, too, we see a form of selectiveness at work where individuals know that there are types of information that they want to find and, more importantly, as articulated by P21 in the opening quote in this chapter, types of information that they do not wish to find. As a result, online search for individuals is a form of computer mediated communication, one where context could better serve to aid both searcher and searched in the access to information. On the part of the searched, there exists an obligation to make available representative content and context; on the part of the searcher, there exists an obligation of restraint and intentionality with respect to privacy violation.

To validate and augment the results of our interviews, we conducted an online survey on a popular survey site asking participants about both appropriate search practices and

their own search practices. As well, we explore more deeply how the relationship between searched and searcher in the online domain influences the relevance of specific types of information. Finally, we explore agency in online search, showing that it can enhance the perception of control for individuals being searched.

The result of our interviews and survey is a design sketch [73] that provides for context and agency in search results. Note that, by design sketch, we are not implying low fidelity prototype. Instead, as noted by Buxton in his book on Sketching user Interfaces, our goal is to suggest and explore the potential realization of agency in user interface [73].

The remainder of this chapter is structured as follows. First, we explore related work in online reputation. Next, we present our interview study, including the design of the study and the results gleaned from a grounded theory analysis. We then present the design of our online survey and the results. Finally, we conclude with a discussion and design sketch which actualize the results of our qualitative inquiry.

2.2 Related Work

Individuals' practices managing and disclosing information online is a frequently explored topic. Social network services provide users with an ability to share, but one burden of social networking is access control [124, 4, 235]. Researchers have noted an increasing awareness of privacy settings over time [25, 259, 244]. However, despite increased awareness, most still tend to use the default settings [243, 235], and Labitzke, Taranu, and Hartenstein [258] showed that only between 7 and 22% of users conceal all possible information from strangers. While researchers accept the fact that over-disclosure is not desirable, Conti and Sobiesk surveyed college undergraduates and showed that students believed that "an honest man has nothing to fear" [106], i.e. that there were limited downsides to information disclosure.

Alongside privacy settings, there exist a number of tools that can secure information. Passwords, and particularly good password practices, are an obvious line of defense against information disclosure, yet compromised data sets of passwords argue that users do not consider security when choosing passwords [156, 330]. As well, many password policies enforce poor password practices, either via a mistaken analysis of relative security or by failing to recognize the mental burden placed upon users who must remember multiple unique passwords [7]. In a related vein, software tools exist to monitor systems including virus scanners and anti-malware/anti-spyware. While good security practices are always advisable, there are still many situations where security practices alone are insufficient to protect a user's privacy or guard against the presence of undesirable content online [511].

Why should one care about online information? The primary answer to this question is the well-established fact that people do search for others online [191]. In an analysis of searching for others, Weerkamp et al. [493] used search engine logs and identified three types of people related search queries based on how well-known a person is. If a person being searched is considered 'low-profile', Weerkamp et al. reported that 38% of these queries were searches for friends, 46% were for people from a searcher's past, and 31% were a search for colleagues. Furthermore, surveys of business practices [37, 306, 104, 384] and anecdotes from interviews indicate that businesses do perform online searches during the vetting process for new employees.

One significant risk of online information is that an online search targeting a person may find information that is perceived to be damaging or unflattering. Even if not damaging or unflattering, a searcher may find information that the searched individual does not wish them to find. The importance of controlling and managing one's online persona is recognized as an important practice, both to control liability and to realize benefit from online information: Many books exist on reputation management [37, 306], and a number of companies exist which offer reputation management services for individuals [104, 384]. Search engines, while a tool that can serve information that is unflattering or damaging, are also a tool frequently used to monitor online content about an individual through self-searching or vanity search: Marshall and Lindley [300] examined and identify motivations for self-search or vanity search, including maintaining online identity and discovering reactions to on-line comments and postings made by the user.

Second, beyond persona management, the more extensive a data set on an individual, the less privacy that individual has. Online stalking and victimization [511] can occur when individuals are too accessible online. As well, when larger data stores are created, the risks to privacy become even more acute [393, 332].

Finally, there remains the question of how best to eliminate content once it appears online, and here the options available to individuals are less rich. The European Court of Justice's Right to be Forgotten judgement requires search engines to comply with requests to remove content about an individual [14, 450]. As well, individuals can request the removal of information, allow information to age or try to respond in some way, but these strategies are not always successful [511, 503].

2.3 Research Overview

Our initial research in this area was motivated by research in reputation management [37, 306, 511] and research in people-related searching [300, 493, 191]. We were also motivated

by stories recounted to us of employers commenting on online content in an interview or even, in one egregious case, asking an interviewee to log in and show her social networking account page during an interview. We began by asking ourselves what perceived limits on appropriate behavior were in the internet era, i.e. what were the social norms around searching for others. We also wanted to explore both if and how people delineate between what they want to share, with whom they want to share information, and what behaviors arise as a function of universal access and persistence of online content.

To answer these initial questions, we conducted an interview study of 21 participants, all of whom had recently been job seekers within the local community. We looked at how these individuals engaged in persona management in the online domain. Given these data, we synthesized a series of interim hypotheses (or theories) [107], designed an online survey and captured data from a sample of 128 participants recruited via an online survey site to triangulate our initial data and to extend our results beyond a single demographic (recent job seekers). The next two sections explore each of these studies in turn.

2.4 Interview study

2.4.1 Study Design

For our initial interview study, we recruited people from our local community who had experience in applying for full time or internship positions. We collected data from 21 participants ranging in ages from 20-47 years of age. Among them were undergraduate students, graduate students, and employees and hiring managers from local industry. Within this group of participants, we performed selective sampling seeking saturation in our data. We began with undergraduates in our co-operative education program. As saturation was reached, we expanded our demographic to include recent graduates, graduate students who had job seeking experience, and more mature employees from local industry.

The study was conducted in-person with each participant. Sessions were approximately 45 minutes. At the beginning of the session, each participant was asked to fill out a questionnaire. The questionnaire at the beginning of each session served two purposes: first, to collect quantitative data about participants' experiences and behaviors online, second, to prim participants, encouraging them to start thinking about their management of online content. The questionnaire included 13 multiple-choice and 3 open-ended questions on personal websites, public blogs and news websites, social networks and searching for oneself and for others. After participants completed the questionnaire, we conducted a semi-structured interview exploring participant's online content management practices and the

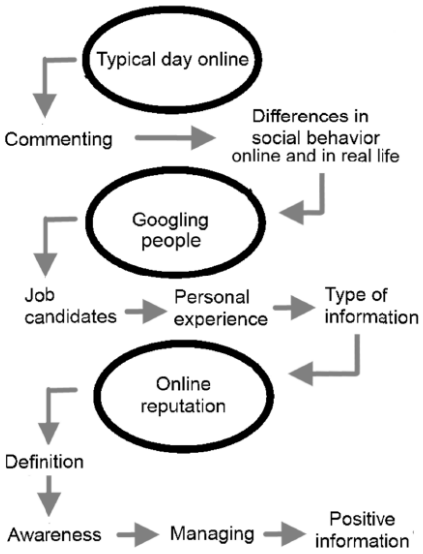

Figure 2.1: Themes explored during the interviews

ways they formed perceptions of others via online content. We chose a semi-structured interview to probe for stories and speculations around topics during the interview. Fig. 2.1 depicts the themes explored, in sequence, during our semi-structured interview. With permission, interviews were audio recorded and then transcribed for analysis.

The data was coded on the ongoing basis during the study to assess when saturation occurred. Coding and theme building was performed collaboratively by the researchers involved in the data gathering.

2.4.2 Result

Questionnaire data allowed us an overview of our participants' online activities. Our participants spent, on average, 6 hours per day online. Most started their typical day visiting news websites. An overview of these online activities of our participant population can be found in Fig. 2.2. Of note, our participants spent significantly more time online reading content than posting content, with most of our participants claiming they never or rarely post content.

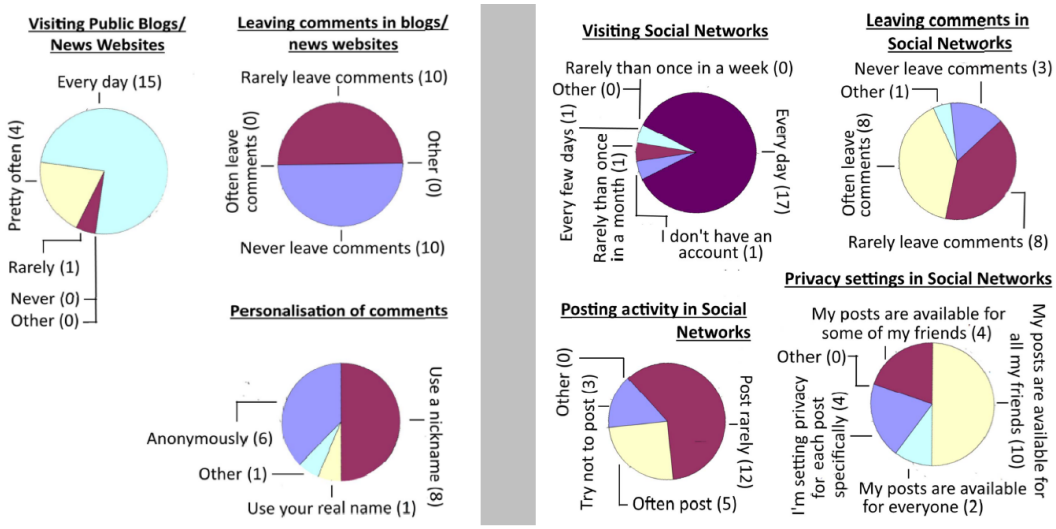

Figure 2.2: Left: Reading and commenting on public blogs and news websites. Right: Social networks visiting, commenting, posting, and use of privacy settings

Selectiveness in one's online presence

Despite limiting activity online, over time content generated by an individual does appear in the online domain and all of our participants had created some personal content online. When we probed our participants about the online content that they have created, we found that two main themes emerged. The first involved controversial content, and other information that might engender mixed reactions from others online, i.e. agreement and disagreement. In the case of these data, participants were careful to conceal their identity, either through anonymity or the use of nicknames. Both our questionnaire (Fig. 2.2) and interview data describe this behaviour:

> *P16: "Almost never associated with my real name... I understand how online interactions are tracked. I have a lot of opinions and they change over time so I really don't want my name to be associated with something taken out of context in the future or even today".*

Another behaviour we identified by our participants was a suppression of content that was perceived to be overly personal. With the exception of one participant, P1, our participants were all frequent users of social networking sites, but described limitations on their behaviour on these sites. In our questionnaires, this emerged in a distinction between visiting (17 visit every day) and posting and commenting (over half either 'rarely' or 'never' post and comment). In our interviews, participants also noted the limits on the

distribution of personal content:

P3: "Some people always share opinions more publicly, or tweet a lot, or they post a lot of photos on Facebook - I tend not to because I don't want to share all of my information with everyone".

While on-line content provided some liability, it was also the case that our participants recognized that avoiding this on-line presence was also counter-productive.

P4: "If you can't find [someone], they are probably trying to hide something. You should be visible"

That information naturally led us to the question of what types of content in general people would and would not like to see about themselves online. At a superficial level, the answer to this question was that job-related information and skills was useful to have online, whereas personal information (opinions, photos, habits) was something participants did not want appearing online. However, with probing, many of our participants described a need to share personal information beyond simply work history and skills. The goal of placing content online was to allow searchers to develop an impression of what it was like to interact with the person they were searching.

P2: "Pictures of travelling, having good time, it is almost like a lie. People do it. It looks good"; "Showing social popularity, showing that you are a nice person to be around"; "What type of connection you have, what type of people you hang out with. Just to judge whom you value as people"

Essentially, one benefit of the online world is that it promotes communication between individuals. When a user searches someone online, the content generated to represent the individual represents the opening lines of a discourse between target and searcher.

2.4.3 Expectations Regarding Others' Online Presence

One goal of our study was to explore norms associated with searching individuals on-line. Because our initial interviews were focused on a professional context, our interviews began by exploring socio-cultural norms around performing web searches of others in a professional context.

We began this section of the interview by asking participants to imagine that they are about to hire a new employee and have several good candidates. All but one of our 20 participants said they would perform an online search for candidates. The claimed motivation for their search would be to verify professional information.

P9: "I want to see [...] what cool things have they [have] done in the past, both professionally and casually"; "But I wouldn't go too far. Usually their professional website will pop-up and I just go there. I'm not looking for anything deeper, I'm just curious to learn more about their work".

However, on probing, it became clear that participants did want more than just a resume to appear online:

P10: "I want them to show the passion about their job and show that they have other interests as well; that would be the ideal case".

The action of searching for someone online was something to be expected by our participants. In participants' provision of content, their expectation of content on others, and their willingness to search for content on others all argue that a simple web query on an individual is appropriate.

Selectiveness: Calibration, Tolerance, and Limits

Given our interest in reputation and the expectation that information exists online, one question we had was how pervasive search was in capturing information about another person. Do individuals simply follow the first links they find? Are they looking for negative or positive information and does there exist a bias toward types of information and the impression created? What are the socio-cultural norms around search? Is anything fair game or is there a limit on what should be explored? Essentially, we wished to explore the concept of selectiveness with respect to their interpretation of content, their tolerance for negative content, and the limits or boundaries of the information they wished to find.

First, our participants were consistent in their articulation of a willingness to calibrate the information desired to the position they were "hiring" for:

P7: "It depends on a position. If the position has to deal with people then I really care about problems this person has had with people before. If I see good pictures, that he has a lot of friends, for example, so he can deal with people so this is a good sign. If this position does not need to talk to people then doesn't matter, all what you need to do then is to make sure that he is not going to kill someone."

Second, we explored the implications of undesirable content that a searcher might encounter online. Undesirable content exists in many forms: content posted by the individual that creates a negative impression (e.g. photos, comments, blogs); social content posted by others that can be deemed embarrassing (tweets, photos, stories, profile on a dating site);

and professional/personal content online that exists because of wrong-doing (e.g. petty crimes, academic dishonesty, etc.). Typically, our participants fell between extremes: they acknowledged that, if they search an individual online, then not all content they find will be positive, but they also do not want the initial impression they glean to be negative.

P9: "I don't mind [if] drunk pictures exist but if I google someone I don't want it to come up, at least on the top".

P1: "If I see extremely shiny reputation that would make me suspicious".

One feature that was of significance in calibration was time. Participants frequently expressed the idea that individuals can change, and that old information is less relevant.

P14: "If it happened a while ago then it's fine"

However, even then, the temporal saliency of information was influenced by context – the age of the other person, their role, the relationship between the searched individual and the searcher. P7 summarizes this point with respect to age:

P7: "Sometimes in life people are less judgmental about some mistakes you've made... over time. So if you are young and you get drunk this might seem almost fine but if you are managing a company and you get drunk I think this is irresponsible. ... I think the older you get the less tolerant people are".

Third, beyond tolerance to negative content, we invested some time exploring the socio-cultural norms around search for others. While our participants did acknowledge that it was expected that people would search for others online, our participants also felt that finding personal information while googling others was a violation of privacy, and that limiting the exploration of content was advisable in their search of others.

P10: "I'd rather find out more personal things organically, in conversation, when they bring it up first [rather than via online search]".

P20: "I'd google someone in a professional sense but I wouldn't look at their Facebook profile. Because I don't think it is anyone's business".

2.4.4 Discussion

As searchers, our participants were willing to: (1) Prioritize content most relevant to the context of search; (2) Interpret content with respect to social context and time; (3) Accept limits on the types of information they should seek online. In synthesizing our data and

building theory, two primary themes have emerged: agency and context. Unifying these attributes of agency and context was the principle of selectiveness [131]. Selectiveness was perceived as desirable both when being searched and when searching others.

Agency, or the desire for some control over what others find online, was a frequent theme in our interviews. The individuals that we interviewed perceived that an online presence was required, and they wanted to ensure that certain content existed and was found by potential searchers in any search of their profile. Currently, online tools only indirectly support agency. One can create a personal web page or a profile on a professional networking site. However, the prominence of this information (and, indeed, whether or not this information is found at all) is subject to the vagaries of ranking algorithms during online search. One frustration that did exist for our participants was their inability to highlight particularly salient information that might be desirable for a potential searcher to know (e.g. professionally, any successful projects, any significant awards; personally or socially, any prominent roles in community organizations). One can highlight this information on a resume, but there is a perceived cachet associated with information being a prominent component of one's online profile.

Alongside agency, a second theme in our data was context. In our initial interviews, we focused primarily on job seekers, and asked these job seekers to explore online content both from their perspective as an employees and their perspective as a hiring manager. In our biased sample, one thing we did note is that, because the focus was on seeking employment and identifying employees, that context heavily influenced the type of content individuals wanted to find. Participants actively noted that the professional relationship implied by hiring meant that professional information was what one should be looking for online. Furthermore, with respect to context, time was a significant factor for many of our participants in the relevance of the information. There was a perception that fresh content was salient, and that aged content less so, particularly given the professional nature of the relationship being explored. One interesting question is whether temporal context plays as significant of a role in saliency when other relationships are at play, e.g. when searching a potential romantic partner online.

Agency and context are related to the issue of selectiveness. Participants wanted to ensure both that desirable information was appropriately ranked and found and that information that was inappropriate to context could be avoided. This issue of selectiveness arose in both being searched (e.g. by a potential employer) and in searching (e.g. as a potential employer). When searching for others, selectiveness was desirable because of a perception of appropriateness: It was perceived of as reasonable to explore professional information, but not personal information. In some participants, there was awareness that legal requirements (e.g. human rights legislation) limited the type of information that

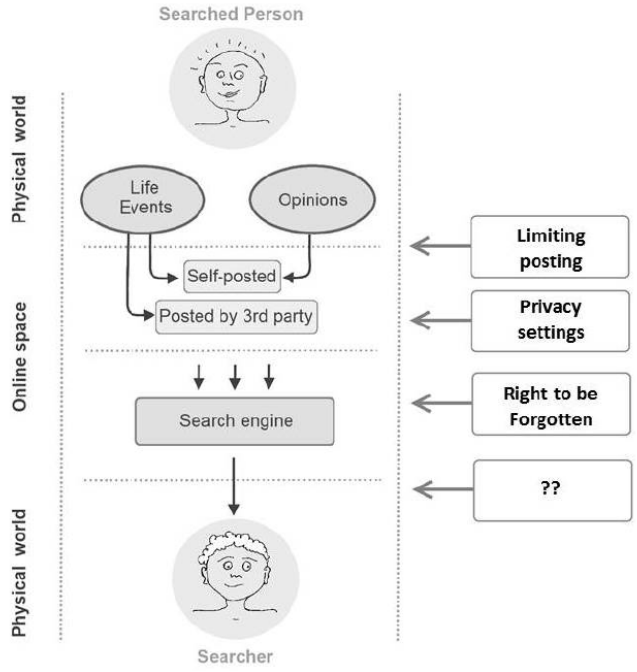

Figure 2.3: Exploring designs supporting agency, context, and selectiveness

one could know and act on in hiring decisions, and that even discovering certain types of information accidentally could disrupt the perception of judgment in decision making.

Search engines, in their indiscriminate provision of information, do little to address selectiveness. As shown in Fig. 2.3, a searched person can limit posting, manipulate privacy settings, and occasionally suppress content either through legal remedies or through search engine optimization techniques (e.g. astroturfing). However, this is indiscriminate selectiveness. It does little to address online content posted by third parties. It does little to provide context in search. It does little to support desired selectiveness on the part of the person seeking information online, i.e. the searcher. Given our observations, it appeared that there may exist a design opportunity within the presentation of search results.

One concern in any discrete study is the risk of overgeneralization. In our interview study, our sample primarily included white-collar job seekers with university-level education within the local community. Furthermore, in the interview setting, there may be a reticence on the part of the interviewee to be completely honest about social norms around online search. In the real world, individuals can peek in windows, stalk people in their daily lives, and in other ways invade privacy, but the ability to monitor actions in the

physical world may inhibit this behaviour in a way that the online world does not. As a result, the observations that motivate the above hypotheses may be biased by sample and context, both of which may color the answers participants provide. To both broaden our demographic and to contextualize our closer to the anonymity of the online world, e conducted a follow-up online survey to test the above hypotheses and to expand on the information obtained via our interviews.

2.5 Online survey

2.5.1 Study design

Given our desire to validate the themes of context, agency, and the principle of selectiveness in online search, we designed our online survey to both verify the concept of selectiveness in information searches and to explore facets of selectiveness: its nature, boarders, granularity, and implementability. Fig. 2.4 depicts the survey structure. Based on previous research [234, 340, 103, 330] and our interviews, we hypothesized two possible bases for selectiveness - social relationship and time (or age) of information, e.g. social context and temporal context, and our online survey focuses on these facets.

Survey participants were residents in the United States. To ensure geographic dispersion we conducted the survey in two phases. First, we captured data from 55 participants between 5:30 – 6:30 pm EDT. We also relaunched the survey between 10:15 - 11:00 pm EDT (7:15 - 8:00 pm PDT) and collected responses from additional 64 participants. Overall, 119 people (57% female) participated in our survey. Based upon survey participants demographics, the median age of our participants was approximately 35-40 years or age.

2.5.2 Data Analysis and Results

We separate our survey data into two sections. The first explores behaviours of our participants when searching others. The second section explores their perceptions of online information about themselves.

Searching Others

Our survey began by asking "How often do you Google a casual acquaintance or someone you have just met?" Surprisingly, 43 of our participants (36.1%) indicated that they

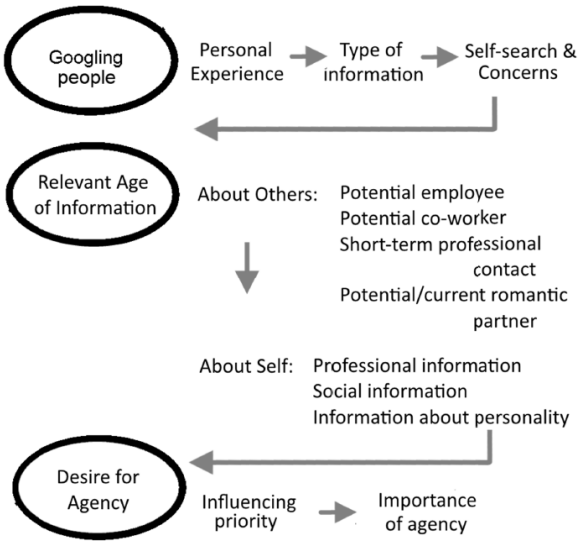

Figure 2.4: The structure of online survey

had never searched for an acquaintance online. Because a significant number of our questions asked about online searches of other individuals, we partitioned our data into two categories: those who had searched for personal acquaintances, and those who had not.

While our demographic data on ages did not permit statistical analysis, qualitatively we do note that those participants who had never searched for an acquaintance online skewed older: 46% of participants who had never searched for an acquaintance were over the age of 45 versus 29% of respondents who had searched for an acquaintance.

We next queried participants on what type of information they would seek on a casual acquaintance. Discriminating between those who had searched for an acquaintance online and those who had not, we found that professional background was the most common reason to search for an acquaintance and there was no statistically significant difference between groups. We also asked about social information (e.g. hobbies and activities), personal information (e.g. opinions, interests, preferences), and whether physical attraction would motivate search and found statistically significant differences between the two groups (Fisher's Exact Test, p-values ¡ 0.0001, 0.0001, 0.05 respectively for social/personal/romantic interest).

Finally, we asked individuals about the length of time information remained relevant in each of these categories with respect to an acquaintance (Fig. 2.5, left). Qualitatively, we note the following: participants who had not searched for an acquaintance generally

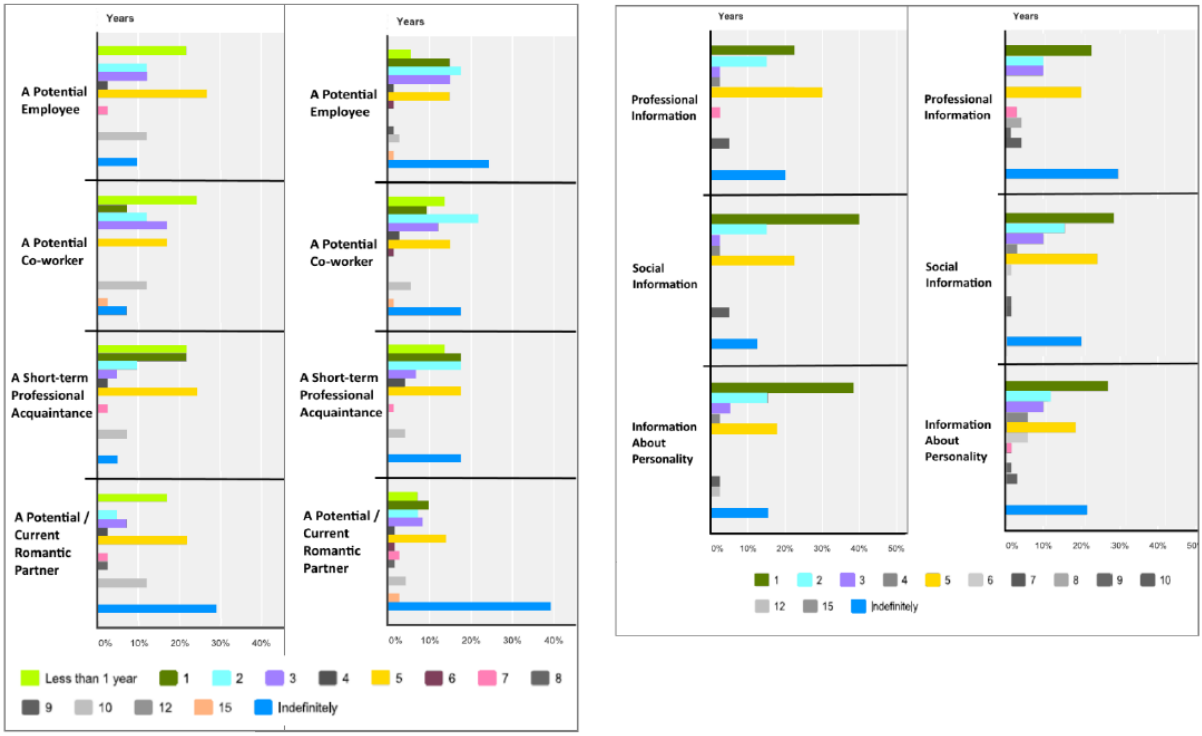

Figure 2.5: Left: Relevant age of content for those who had never searched someone (left) and who had (right). Right: Relevant age of content on self for those who had never searched someone (left) and who had (right).

had shorter time frames for the age of information that was relevant (see spike in less than one year data on the graph and corresponding reduction in indefinitely). Furthermore, for both participant groups we see a cluster of data less than five years and then longer periods typically defaulting to indefinitely. Finally, the relevance of information is skewed longer for a romantic partner, with a significant spike in information being relevant indefinitely.

Searching One's Self

Our second area of inquiry explored awareness of information about one's self online. First, we looked at self-search of vanity-search, and asked participants about awareness of results associated with their name and concern about search results associated with their name. We see a statistically significant difference between the two groups of participants with respect to awareness of search results on their own name (Fisher's Exact Test, p ¡ 0.001)

but no differences in concern about online information (Fisher's Exact Test, p ¿ 0.10).

Next, we replicated a question on online search on others and asked participants about the age of interest for information about themselves. Interestingly, one thing that becomes apparent in this result is that, particularly for people who had searched for others online, the temporal relevance of information was shorter (Fig. 2.5, right).

Finally, we asked participants about agency in the presentation of online search. Specifically, we asked the following: *If you had an option to pin one or two top-level search results on your name, would you?* This feature was particularly important to those who had searched for information on others online. 76% of participants who had searched online for others stated that they would make use of this feature (versus 30% of those who had not).

2.5.3 Discussion

Three implications arise from our survey data. First, we find that temporal context is important in search results. Second, we find that social context also matters with respect to the importance of search results. Finally, we find a desire for agency in the presentation of search results on an individual, particularly for those participants (approximately 2/3 in our survey) who had searched for acquaintances online. Unfortunately, the modern search infrastructure provides no support for any of these attributes of search on an individual. The user of a search engine is perceived to be the searcher, not the searched.

2.6 Implications for Design

In synthesizing the results of our interviews and surveys, it appears that there are significant opportunities for the redesign of modern search engines to support both searcher and searched as users or communicators in a computer mediated process. We synthesize this design around the concept of *Selective Sharing* and note that both social and temporal contexts influence the saliency of information presented. Fig. 2.6 depicts the principle of selective sharing as gleaned from data collected in our interviews and in our online survey.

Expanding on Fig. 2.6, first, from the social context perspective we argue for the separation of "professional", "social", and "personal" online personas [340]. In our interviews, this appears to be a common delineation between categories of search results. Furthermore, the categories have sufficient granularity that tractable semantic analysis should be able to quasi-automatically segment content for users, potentially with some agency for users [42] to refine the labeling of relevant social context via vanity search.

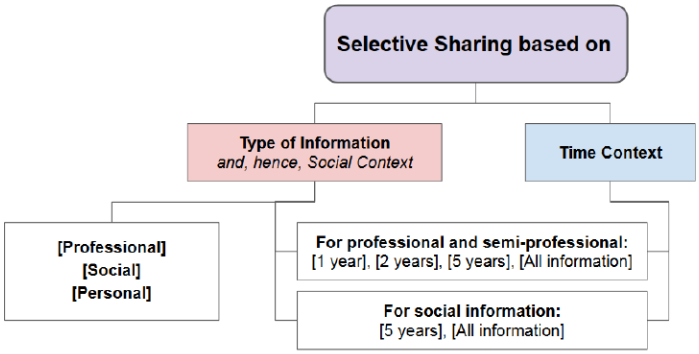

Figure 2.6: Analyzing factors associated with selective sharing.

Second, we suggest that the following time periods are of interest to the searcher:

- 1 year, 2 years, 5 years and indefinite (all information) for professional and semiprofessional personas;
- 5 years and indefinite (all information) for social and personal information.

In synthesizing data from our participants and information from past research, we believe that online search providers can promote selectiveness. We provide a "sketch" [73] of a "personal search page" (Fig.2.7), where search results on an individual are provided.

In this design sketch, selectiveness, agency, and context are supported in multiple ways. When searching an individual, the top search results for that individual (shown immediately below the search bar) represent pinned content, or content prioritized in some way by an individual. Below that, different facets or social contexts represented in search results are shown and can be highlighted. As well, the temporal context of search results is explicitly supported via arrangement on a timeline. Searchers can still access all information on an individual; however, each result is contextualized both with respect to facet and time. As an analogue to real-world privacy invasion, it is true that individuals can still choose to peek into someone else's windows. However, this search interfaces also allows a searcher to *choose not to* explore irrelevant content.

How do the concepts in this sketch support selectiveness? First, in highlighting and pinning content, a searched individual is selecting specific content as of primary place and importance, i.e. pinned content supports *agency*. Search providers already support personalization in the presentation of search results via user accounts. Undoubtedly, this content may change with time and phase of life: a job seeker may have different ideas of what pinned content should be versus a retiree. Agency allows this customization

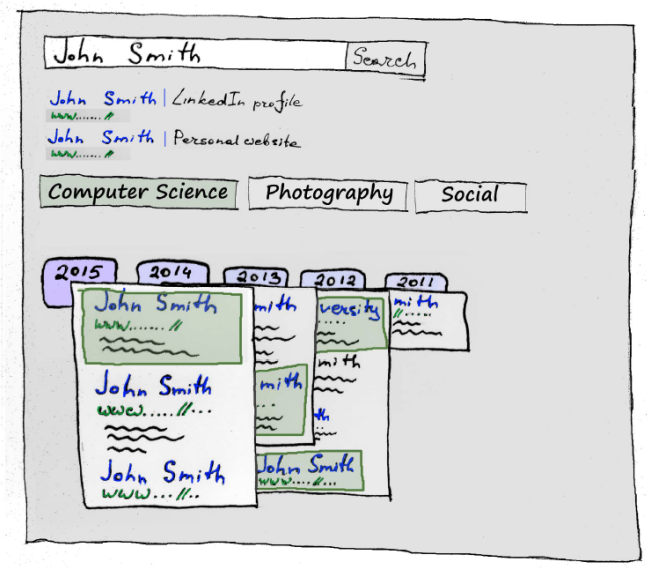

Figure 2.7: An early, evocative design sketch to support discourse around selectiveness, agency, and contextualization of information.

for time and phase of life. Second, in displaying pinned content and highlighting the social context of information, searchers can practice selectiveness by only following pinned content or content that is highlighted and characterized in a specific way. As well, via a timeline searchers can also practice selectiveness by only selecting content of a certain age. Finally, even if a searcher decides to look deeply into all content, searchers are forced to, at minimum, contextualize the content as being about another aspect of an individual's life or from a certain specific time frame of a person's life. Given the resources available to a modern search engine [42], including timestamps, keywords, locales, data triangulation, and personal accounts, we posit that it is feasible to explore search interface designs.

2.6.1 Limits and Benefits of the Design Sketch

As we noted earlier, we do not view this search result as a low-fidelity prototype or as a finished design product [73]. Often, when presented with a design sketch, the first objection of a reader is that it does not solve a specific problem [73], a plight that has been well highlighted by Greenberg and Buxton in their essay on usability evaluation [184]. For example, how does the design sketch help prevent someone who is intent on finding negative content from finding that negative content if it exists? It does not. How does

the design sketch allow an individual to remove content that they wish was not available online? It does not. However, it is vacuously true that there is not one silver bullet that will address every single instance of privacy violation online. As one trivial example of this, using privacy settings is only foolproof if website security never fails and passwords are never compromised.

The internet age's new norm is instantaneous access to and indefinite availability of every single piece of information ever posted about an individual. Unless we change this attribute of the online world by allowing the censoring content in some way, privacy must become a social contract between searched individuals and searchers. In our data, searched individuals and searchers both articulate a willingness to practice selectiveness in their choices. The purpose of this design sketch is to begin a discussion around this principle of selectiveness and how best to support it for willing participants.

Alongside the onus on searchers to practice selectiveness, one added onus exists when considering the principle of selectiveness in online information as articulated by our design sketch. While supporting voluntary selectiveness in online search is technically feasible, search providers must be willing to take steps to aid individuals in adopting selectiveness. On one hand, this onus requires effort on the part of search providers, which may, at first blush, appear unfair. On the other hand, the Right to be Forgotten ruling in Europe is one example of the alternative faced by search providers. The principle behind the European Court of Justice ruling is that providing content results in an obligation to, at some level, also vet that content, particularly if a party feels that the content is a misrepresentation. The challenge is to satisfy the vetting of online content in a measured and appropriate way. Do we suppress content, as in the right to be forgotten ruling? Or do designs exist that can help to mitigate of contextual negative content in a way that eliminates the need for censorship of results by search providers?

2.7 Conclusion

This work describes our exploration of online content, online persona, and online reputation management. Results of an in-person questionnaire and qualitative interview study (n = 21) and an online survey (n = 119) indicate that participants vet content they post online. As well, participants consider context, specifically both social context and temporal context regarding online material. Participants indicate that they believe that limits exist on how aggressively they should seek content on others online. Finally, participants indicate a desire for agency via an ability to prioritize and categorize content.

We link these concepts together via the higher-level principle of selectiveness [131], i.e. the desire of both a searched individual and a searcher to constrain the information to the specific goal of the search. To advance our discussion of the concepts of socio-selectiveness, temporal selectiveness, and agency, we synthesize an evocative rough design sketch [73] which can serve to ground the discussion of socio-cultural norms and practices and how they can best be supported in online search.

One aspect of this work is that the connection between individuals may vary along a spectrum of intimacy. However, the relationship between searcher and searched is rarely as directly intimate as relationships within, for example, family units. In the next chapter, we look at on-line posting behavior of parents of children, i.e. sharenting practices, with the goal of understanding how posters conceptualize privacy and agency when they are making decisions for their child.

Chapter 3

Project 2: Perceived Protagonist of Information Shared Online

In this chapter, to explore online posting of information that is more intimate (and information that the poster would not want misused), we explore the posting behaviors of parents. In particular, we examine sharenting posts to understand whether the intention is to post information about ones-self as a parent or about a child. We also explore how perceptions of who the post is about does or does not impact posting behavior.

3.1 Introduction

Almost a decade ago it was anticipated that millennials would retain their willingness to share personal information online even as they get older and take on work and family responsibilities [16]. Their tendency to perceive advantages of personal disclosure [452, 521] was predicted to outweigh their privacy concerns. As millenials matured, they continued [439, 318] to actively engage in sharing information about their experiences, including experiences as parents. The phenomenon of sharing information about one's child online became so popular [64, 487], that is now labeled as a distinct social media behavior – sharenting (from "share" and "parenting") [55, 64, 444, 98].

Arguably, the information posted by parents over time becomes a component of a child's online identity. Thus, sharenting practices result in a child's digital identity being formed from a very young age [64, 444], before children gain an ability to express their own agency

[114, 98]. The emotional reactions of teenagers and adults suggest that issues can be expected for children growing up with an online identity that includes prior content posted by others [114]. For instance, research indicates that adolescents are often embarrassed about the content share by their parents [167]. In fact, many parents have already been asked by their child to remove some of the information shared about them [148]. However, despite parents' awareness of the social implications of digital parenting [87], parents continue to post revealing information about their children [98, 64]. These practices seem to be particularly persistent due to the variety of advantages that parents gain through sharenting, e.g., connecting with family and friends [64], creating an identity as a parent [102, 114, 246], archiving memories [55], and receiving social support [308].

This realization inevitably brings a complex problem of balance between the parent and child identity representations through the information disclosed online [55, 487, 64, 444]. While parents generally acknowledge that they have to balance between their child's privacy protection [246, 487] and their own rights to gain advantages from sharenting [487], the functional nature of this balance is poorly understood [55], i.e., what does it actually mean to maintain a proper balance between identity representations? The strategies, currently adopted by parents, mainly revolve around stewarding the child's privacy [246] by identifying appropriate information [15, 246, 55], controlling information dynamics [15, 64], and limiting oversharing [15, 87]. However, these are ad hoc strategies, predominantly focused on posting decisions and activities, which place the burden of data management mainly on parents [15].

The self-presentation aspect of sharenting occurs through a form of a storytelling using posts on social media [497]. In any form of a social story, the central character is referred to as a *protagonist*. Since parents create their story on social media as a means for self-presentation [102, 114, 246], one could assume that they would naturally become a protagonist in these stories. However, when a story involves other identities, in this case one's child, the identity of the protagonist may become more nuanced. Then, where does a parent end and a child start in a sharenting post? And can we, as technology designers, learn to recognize this context to provide a required identity separation to assist in longitudinal digital identity development?

Motivated by the conflicting co-existence of the parental self-presentation benefits from sharenting [102, 114, 246] with the child's later self-identification with the shared information [167, 148], the above questions highlight the fact that the perceived identity associated with the sharenting information may be ambiguous and context-dependent. While research demonstrates some aspects of this ambiguity – e.g. the co-existence of self- and child-directed motivations in posting activities [246] – and discusses the consequences of dual identity representation – e.g. social [114, 167], legal [444], and ethical [55, 487] chal-

lenges – neither the explicit existence of protagonist ambiguity, nor the underlying nature of the designation of a protagonist have been previously discussed. Yet, if the protagonist assignment in sharenting posts is indeed ambiguous and context-dependent, this understanding would open important development directions for systems design, online privacy support, and digital identity management. For instance, to appropriately support digital identity management for both current and future users (parents and children respectively) [511, 250], recognizing about whom this information is perceived to be in a given context can be leveraged to correctly filter and represent this information [250].

Through a set of semi-structured interviews with parents of young children (n=23), we explore self-reported factors taken into a consideration when assigning a protagonist to a sharenting post in two contexts – as a sharing parent (an author), and as a viewer of a post. Consequently, our analyses present two inter-dependent research contributions. First, we demonstrate the presence of the ambiguity of a protagonistic identity for the same post material, and discuss what contributes to this tension of the protagonist designation (RC1). In particular, we demonstrate that the same post is perceived to portray a child when considered from the perspective of potential safety issues or the future information perception. However, when the same post is considered from the perspective of communication between a parent and their expected audience, the protagonist tends to gravitate towards the parent. Second, we explore the mechanisms associated with assigning protagonist identity to sharenting posts (RC2). We identify four factors affecting the designation of a protagonist: locus of sharing motivation, agency in sharing, social connection to the information recipient, and agency of the captured action. We continue by analyzing sharenting posts through the lens of a trasactional communication process [34] with "imagined" audiences [278] We demonstrate a conflicting co-existence of two sharenting models – for "intended" and for "unintended" information consumption. We discuss how the discrepancies between these two models relate to protagonist ambiguity and to the broadly discussed tension between the parental benefits from sharenting and the child's privacy.

The remainder of this chapter is organized as follows. We first provide an overview of the extensive body of related work on sharenting. We then present our study design and analysis of our results around two research contributions. Finally, we highlight our consolidated analyses and their implications for tools to infer protagonist in posts.

3.2 Related Work

The practice of parents sharing information about their children online, often referred to as sharenting [55, 64, 444, 98], presents a number of challenges from both social, ethical,

and technical perspectives. For example, Steinberg [444] analyses the problem of parents posting information without a child's consent from a legal perspective, highlighting the dual role of parents in their child's identity both as "gatekeepers" of personal information, and as creators of their child's life story. Blum-Ross and Livingstone [55] discuss ethical dilemmas of blogging parents, and show that, while parents are responsible for safeguarding their children against privacy risks, representing parental identities online inevitably exposes aspects of their children's lives publicly [55]. Similarly, Wagner and Gasche [487] demonstrate the controversy of sharenting through the fact that parents have to balance their children's privacy interests with their own right to express themselves online.

Parents generally demonstrate an awareness of the need for this balance [87, 55], and engage in stewarding the child's privacy [246] by identifying appropriate information [15, 246, 55], controlling information dynamics [15, 64], and leveraging affordances of social network services to limit oversharing [15, 87]. However, this balance between the parent's identity representation and the child privacy rights in sharenting remains challenging [98, 64, 246], especially since sharenting practices bring significant social benefits to parents. Such benefits include helping parents to connect with family and friends [64, 246], supporting the development of their self-presentation as a parent [102, 114, 246], or providing means for receiving validation, a sense of solidarity, and social support [308, 287, 246, 87]. For example, Kumar et al. [246] note that while some mothers are concerned with controlling digital footprints and oversharing, the benefits of receiving validation might outweigh their concerns. Similarly, mothers were shown to reveal personal information about their children to participate in social interactions and to show their pride [487]. These social benefits of sharentning naturally extend the variety of social advantages that people gain through personal information sharing practices in general. For example, sharing of personal photos was shown to serve many interpersonal functions, such as forming new relationships and maintaining old ones [163, 476], and narrating people's histories [484, 476].

The social aspect of sharenting might play a particularly important role in information disclosure decisions. For example, research on the factors influencing photo sharing [40] shows that sharing intentions are positively influenced by self-presentation and communication benefits. Taddicken [458] demonstrates that the social relevance of an application is a significant predictor of users' willingness to disclose personal information. Finally, people reporting stronger social ties are more likely to share more photos online [443], and there is a positive relationship between users' "sense of embeddedness" and the number of photos they contribute publicly [336].

Thus, the problem of balancing children's privacy interests with parents' rights is noticeably complicated by the extensive social benefits that sharenting can offer to the parents. Correspondingly, while protecting one's privacy was previously shown to motivate users to

share information selectively or not to share at all [9, 448], many parents continue posting revealing information, including information about their children [98, 64] that might affect their children's lives [114] or might be misused by others [225]. Besides a number of specific safety issues, such as identity theft, resharing pirated information on predatory sites, and revealing information about a child's location and schedule [225, 64, 341], sharenting practices also contribute to the development of the child's digital identity [444, 98]. Specifically, Davidson-Wall [114] discusses the conflicts that might arise between parental impression management and the restricted autonomy necessary for teenagers to develop their own identity during adolescence. Indeed, research with adolescents reveals some contradictions between the image they are trying to construct online and the posts of their parents, which might create embarrassing situations [167]. In order to avoid conflicts, adolescents indicated that parents should respect boundaries regarding types of posts that can be shared, how often they can be shared, and with whom. Interestingly, Moser and colleagues [315] find that parents and children are generally in agreement on how often and how much information parents share about their children on social media. However, there are disagreements about the permission-seeking process, as children believe their parents should ask permission more than parents think they should, and parents believe they should ask for permission more often than they actually do, especially younger parents. Earlier research also shows that many parents who share information online about their child have been asked by their child to remove some of this information [148].

Taken together, prior research shows that sharenting exists within a persistent tension between the need for children's privacy and the parental benefits of social openness. Moreover, while parents adopt strategies to support their child's privacy, the diverse social benefits from sharenting might often outweigh their concerns. On the other hand, the reactions of adolescents demonstrate that with time, the information shared by parents becomes a part of the child's digital identity.

Correspondingly, we suggest that the challenge of identity representation in sharenting is tightly associated with the contextual ambivalence of a protagonist's identity in such posts. To address this challenge, this project explores how people conceptualize these identity representations, and how they dynamically assign a protagonist in sharenting posts.

3.3 Method

We performed a qualitative study that leverages semi-structured interviews to probe parental perceptions of sharenting information. In this section, we describe our participant popu-

lation, study procedure, and our analytical approach.

3.3.1 Participants and Recruitment

In total, we interviewed 23 millenial [125] parents (age 24-37, ave.age 33.22) of 29 young children (age 0-10, ave.age 4.69). The rationale for the age group of participants is, first, that millennials use social media for personal reasons more than any other age group [318]. Furthermore, there is evidence that digital media experiences during childhood influence user's privacy management later in life [279]. Millenials were the first generation to be exposed to the Internet in their teens, when they were shown to be much more likely than parents to give sensitive information [452] both about themselves and their families [471], especially given perceived benefits from information disclosure [521]. The ceiling of the age-interval for children is defined by previously reported results [167] that ten years as the age when children tend to become aware of the fact that they are represented online in a specific way. It is further motivated by the reports on social media use, showing that children start having social media accounts around the age of 10, such that 18% [86] to 20% [18] of 8-11 year olds who go online have a profile, and 11% of children get a social media account when they are younger than 10 [129]. Four of our participants were couples; others were the sole participant in the parent-child relationship.

All our participants had a tertiary degree. A self-identified gender distribution in our sample was represented by 14 mothers and 9 fathers. While a number of previous studies on sharenting focused specifically on mothers (e.g. [246, 87]), or the differences between fathers' and mothers' posting practices (e.g. [15]), we do not distinguish between parents of different gender. Instead, in our analysis we focus on the overarching trends in protagonist assignment. Due to the nature of our research questions, we also included in our sample parents who post often, rarely, or refrain from posting about their children at all. In particular, we were interested in a comparative analysis of underlying motivations and considerations of these parents from the perspective of protagonist assignment.

Participant recruiting was performed via respondent-driven sampling (RDS) [201] – a variation of chain-referral sampling, when study participants serve as references for additional participants. RDS brings an element of sample randomness by separating several referral chains, thus avoiding biases from sampling only from one social circle. Specifically, we limited sampling to 3 participants from the same referral chain.

3.3.2 Study Procedure and Data Analysis

The study was conducted in-person or via video conferencing software, and informed consent was obtained prior to the interview session.

Our interview followed a semi-structured format exploring parents' posting behaviours, their friends' posting behaviours, perceived benefits and concerns of on-line information, overall behavioural and mental aspects of sharing practices, and views on the persistence and potential future life of information. During interviews participants were encouraged to open their social networking feed while discussing their and their friends' sharenting practices. In particular, participants were asked to walk us through recent posts they had made and posts they had received from their friends. We probed motivations and rationale for posting, and who they perceived the post to be about (a protagonist of the post).

We also provided parents with several scenarios and discussed their corresponding perception of a protagonist. Scenarios described a post made by a participant's friend and included a family photo from a vacation, a child's photo from a vacation, a child's photo from a regular day, and a text description of a family event that includes a child's activity. These scenarios were constructed to, first, capture potential differences in perception of photos versus text-based posts, second, in perception of posts based on the included actors, and finally, we aimed to reflect the previously reported results that parents tend to post photos that are cute, funny, with family or friends, and milestones [246]. According to the semi-structured format of the interviews, the described scenarios were used as illustrative discussion points to probe the differences in the protagonist assignment. Thus, participants would often suggest further details for each scenario to highlight their nuanced perception. We analyse these comments alongside with the other data on factors affecting the assignment of a protagonist.

At all stages of the interview, we worked to ensure that the opinions, concerns, and motivations elicited were those of participants and were not biased by the researchers; in particular, we avoided value judgments and allowed participants to highlight similarities with their friends' posting behaviours.

We coded [107] data in 3 phases. First, we performed incremental open coding on our data. Incremental coding was performed by one of the authors, and codes were then discussed and refactored in consultation with additional members of the research team. Next, we performed axial coding, seeking relationships between the lower-level open coding. Saturation occurred after 12 participants; however, we continued to explore nuance and seek validity in our data over the remaining participants in our sample population. It was particularly important in this study, since we were seeking to understand discrepancies in the protagonist assignment and needed to ensure that these discrepancies were representative.

3.4 RC1 Results: Ambiguity Of a Protagonist

Presenting our results, we start by demonstrating the presence of ambiguity in the identity of the protagonist in sharenting posts. One primary theme that arose during axial coding was that the protagonist in sharenting posts switched between child and parent. As per Corbin and Strauss [107], to understand this ambiguity we specifically explored context associated with it, including how content was posted, who could see it, and how it was expected to be perceived. We also explored mitigation strategies employed by parents. Overall, during axial coding, we found that, depending on whether posts are considered from the perspective of current or future risk or from the perspective of communicative goals, this contextual perspective noticeably influenced parents' perceptions of who a post was about, i.e. the protagonist within the post.

In the remainder of this section, we demonstrate how these perspectives impact protagonist assignment both in the context of authoring a post and viewing posts of others.

3.4.1 Potential Challenges Are Focused On A Child

The perceived risks of sharenting are consistent between parents who restrict and who do not restrict their sharing practices. We categorize these risks into two themes: considerations of safety issues and future privacy issues. We show that in both cases, the perceived risks are concentrated on a child, who is seen as a protagonist in the shared information.

Safety considerations

While parents take the responsibility for guarding the child's safety, axial coding reveals that safety considerations are predominantly formulated from the perspective of a child as the main portrayed figure in the information. Thematically, these safety considerations are strongly associated with the lack of control over the potential information recipients:

> *"You never know where your child's photo will end up. I just feel uncomfortable if my child's photo end up somewhere randomly, and I don't know who will be looking at it. I don't know what to expect from these people." [P6]*

This concern about the lack of control cannot be disentangled from uncertainty about recipients' intentions toward or perceptions of the information – how recipients may consume or redistribute it. The unwanted intentions mainly revolve around sexual or otherwise inappropriate use of the child's photos.

"Any photo can be viewed with some intention that I did not anticipate. Feeling that I cannot control that someone might be watching these photos with perverted intentions makes me paranoid." [P15]
"There are all kinds of people there. I really wouldn't want someone creepy looking at my kids." [P23]
"There are people who might not necessarily have right intentions, that might have negative effects." [P22]

Note, that while the actual posting decisions and sense of threat might vary among parents, the perceived challenges are consistent within posting considerations:

"First, only my friends can see photos of my children, my accounts are open only for my friends. Second, I'm trying not to post some semi-naked photos of them so I don't really know how they might affect their safety." [P9]

Future Privacy Considerations

Our analysis also revealed concerns regarding the future perception of the information. For instance, parents appreciate the fact that the perception of privacy is highly personal, and, in the future, their children's perception of privacy might differ from their own. Respectively, there is a concern that children might disagree with parental information sharing choices.

"My concern is that after a certain age I don't want her to complain about any content that might hurt her privacy, any content we uploaded on the social media." [P2]
"I thought about it. Would I like if all my baby pictures were out there? If he says he doesn't like it I'll say 'okay, let's delete it'. But I'm not sure how he might react." [P3]
"If you post some terrible photos of your children, in the future they come to you like 'why did you do that?'." [P9]

We also found a persistent notion across our participants of a potential effect of the shared information on a child's professional life, and their future personal relationships:

"It is not fair to my child. When in 20-30 years they are searching for a job, I'm not sure all children would say 'thank you' to what their parents posted." [P4]
"What if he goes to army, or something, and this information would not allow him to advance in his career." [P5]
"For example, she meets a guy, they start romantic relationship and then he sees some inappropriate photo from her childhood - who knows how they both react." [P8]

Perceived risks regarding safety and the future perception of the information lead to the problem of information appropriateness, extensively discussed in literature ([246, 15, 64]):

"Probably wouldn't post potty training or something like that. Not on public social media. Something that would obviously be embarrassing for an older person." [P1]
"Obviously, naked baby photos. Like 'Mom, don't show those to my boyfriend, that's awkward'. I don't post naked photos. I wouldn't post anything that exposes him." [P3]
"In 20 years when he is a CEO and someone comes saying 'I found your baby photo where you are running around naked' - that might negatively effect him." [P6]

This analysis shows that when a post is considered from the perspective of either potential safety issues or the future information perception, the information is seen as predominantly portraying a child. In the following section, we discuss the differences in a perceived protagonist in the same sharenting post considered from the perspective of communicative goals.

3.4.2 Communicative Goals Are Focused On A Parent

In contrast to the considerations of potential risks, from the communication perspective the protagonist is more nuanced and more often assigned to the parental identity, both in the context of information sharing and information viewing.

Specifically, while our participants acknowledged the challenges of separating identities within sharenting posts, in the context of information sharing (authoring), parents tend to perceive the shared information as mainly about themselves.

"It's usually stories about me. The last time I've posted photos with [my children] was from our trip. Just some nice picture of them. But I still see this information more about myself." [P16]
"I post it, it is more about me as a part of our family." [P4]
"For me, it's about our family when I post. Me as 'here is me and this is my family'. Not about children." [P21]

Furthermore, the occasional discrepancy in the perception of a protagonist by a parent who is sharing the information and a recipient of this information might be seen as misinterpretation of the message and cause discomfort to the sharing parent.

"I guess, more about me. Someone commented lately that they recognized something because they are following my daughter's life through my posts. I was really surprised to hear that. I realized that I didn't mean to represent her. It's more about the family side of my life." [P15]

Some participants articulated that their author perception of a protagonist is context-dependent. One parent who created a separate account for their months-old child described

different interpretations for similar information on these accounts:

"I post on my account if I want to share the story from my perspective, and on his account if from his perspective. It's more about him, his life, where he is the center. On my account it's mostly same baby pictures too, but it's more about me as a parent." [P3]

We also saw that some parents restrict their posting practices according to the ambiguous perception of a protagonist:

"We never post photos where it is just our daughter, only photos of the whole family, and if she is not obviously visible, like a side view, or she is turned with her back. Because it's either mine, or my husband's profile, not our daughter's. Soon she grows up and she is going to have her own opinion about her own profiles. And she will be deciding the way they are." [P8]

We then analyzed the perception of a protagonist in the context of intended information consumption, when parents see sharenting posts of their friends. We refer to it as "intended" consumption because it reflects the information appearing on one's social network as intended by an author of a post. Similar to the context of information sharing, we observed a strong tendency to attribute the protagonist to a posting parent:

"I think, I see it more as about my friends. Them showing their kids, families." [P21]
"It's parents, expressing themselves through children." [P8]
"In my opinion, it is always about a parent. In all posts it goes as about a parent, even if a parent is not in the post, because a parent decides what to show, what to say." [P4]

However, in the context of viewing the sharenting information, the designation of a protagonist in a post appears to be more nuanced and context-dependent:

"And if it is just a photo of only a child and say from home, or some achievement, for example, then it is more about this child. So it's a context of the photo itself." [P13]
"If I know everyone then it is about the family. If it is someone from work, then about just that friend" [P14]
"I would see it as about parents. But realistically, of course, I understand that it is about their children as well." [P7]

3.4.3 Dynamic Assignment Of A Protagonist

Finally, in correspondence to the results on parental sharenting considerations (sec.3.4.1), we see that in the future a perceived protagonist in the same information is expected to change.While some believe that the information they are posting in the present will be

perceived in the future as about both themselves and the child, most parents expect the future perception of present-day information to shift towards the child.

> *"I actually think in the future, when children grow up, it will already be about them, despite the fact that mothers posted it." [P4]*
> *"No, of course, it is going to change. In the future it would be a representation of her [a daughter]." [P23]*
> *"I think, it will be more about the child. Because it's their face. I never actually thought about how our children are going to feel about these photos. But I think, children would like it. Because it's their life." [P13]*

Such dynamic nature of the perceived protagonist leads to the uncertainty regarding the information appropriateness. The notion of information appropriateness is tightly related to the possibility of damaging effects on a child's life, including effects on their privacy and future reputation (sec.3.4.1). We found that the acknowledgement of the future significance of the shared information, and, in particular, the expected shift of the protagonist identity, causes parents to experience confusion and disorientation on what exact information might be potentially harmful to their child in the future.

> *"Things that might significantly affect my child in the future I would not want to post. But you don't always know what might affect them. For example, a photo in a diaper - is it good or bad? Say, my child wants to be a president - would that affect them?" [P14]*
> *"I don't want them to feel offended or embarrassed by something I showed. But this category 'embarrassing' is very unclear, I can still accidentally do something and they would say 'How could you?'." [P5]*
> *"It's hard. In the future, what will be normal, adequate and what will be judged?" [P7]*

To summarize, we found the ambiguous and dynamic nature of the perceived protagonist in sharenting posts – a protagonist in the same information changes over time and between contexts. First, while sharenting is known to present potential harm for parents themselves [98], our results show that parental sharenting considerations (regarding safety and future privacy) revolve around a child's identity. However, when considered within the communicative frame of reference within which the post was intended, the identity of the protagonist shifts noticeably toward the parent. Furthermore, in the context of information viewing, an assignment of a protagonist seem to be more nuanced than in the information sharing (authoring) context. Lastly, in correspondence to considerations of future privacy, the perception of a protagonist in sharenting posts is expected to shift towards the child, and leads to uncertainties about the future appropriateness of information.

3.5 RC2 Results: protagonist designation

The results in the previous section showed both ambiguity and dynamics in perceived protagonist. To explore what defines this ambiguity and dynamics, we investigate protagonist discrepancies between different contexts. In particular, we leveraged our semi-structured interview format to identify self-reported defining mechanisms that motivated protagonist perception. Across our coded data of parental reasoning and comparative perceptions, we identified four factors that impacted protagonist designation: the locus of motivation, agency in sharing decision, social connections to information recipients, and perceived agency of captured actions. We discuss these factors in detail in this section.

3.5.1 Factor 1: Locus of motivation

The first factor we identified as affecting the protagonist assignment is a locus of sharing motivation – whose interests and social needs are intended to be addressed by the post. Specifically, in the context of information sharing, if the locus of parental motivation is inward (e.g. to connect with others, to self-represent as a parent, to seek aid and support), parents perceive the information as about themselves:

> *"Yes, I need other people to see that I have such amazing children. It's like bragging about them." [P5]*
> *"Recently I posted a photo of a project she brought from her daycare. But I still would say that it was about me. I'm a proud mother, it's a way to express my ambitions." [P8]*

Similarly, the more external locus of motivation, such as demonstrating the life and development of a child to others, the more parents' perceive the child as the protagonist:

> *I created his account and post there little snippets of what he is doing. I started his account so that mine would not be so focused on him. Here is his account – if you want to follow him and his updates – follow here. [P3]*

While parents might consider their sharing motivation as external, e.g., keeping family and friends updated, the viewers of such posts predominantly perceive it as internal to the parent, i.e. the sharing parent is perceived to be the protagonist:

> *"It's about parents. Because they usually post their children more for their own self-esteem, to boost their own confidence – to show what a nice family they have. So that other people say what a nice baby, and so on." [P6]*
> *"More about parents. Them being proud - here is my child, who draws pretty pictures, or something like that." [P23]*

Previous research on sharenting activities has described the co-existence of internal and external motivations in posting activities, for example, motivations to build an identity as a mother, but not to portray a child negatively [246]. However, the effect of this motivation on the protagonist assignment of the sharenting posts has not been shown before.

3.5.2 Factor 2: Agency of sharing decision

This factor in protagonist assignment is related to more formal aspects of information management, such as an actual decision to share specific information and a necessity to acquire relevant permission. We refer to this factor as the agency of sharing decision. Our data shows that a sharing decision on its own is commonly perceived as a means of self-representation. Thus, the lack of a child's ability to express their voice in sharing decision reduces their representation in the post, while the parent's agency increases perception of them as a protagonist.

> *"It is mostly about me, essentially. It's my profile, I am responsible for the decision what information to post there, so it's mainly saying something about me." [P7]*
> *"More about myself. My youngest doesn't care yet, and by the time he might, these photos will be long forgotten. My older son might... but no, he still doesn't care yet." [P16]*

This association between sharing decisions and the protagonist's identity also persists in the context of information consumption, when the post is decoded by a viewer:

> *"It's mostly about parents. It's their initiative. Children don't post themselves..." [P23]*
> *"Whatever person is posting on their own - it's about them. How they want to be perceived by others. While a parent is posting it is about a parent, it's their actions. When a child gets older, they can say something on that. It is about the initiative. Of a parent and then of a child." [P4]*

The decision to share information seems to be tightly intertwined with the perceived implied agreement for the information to exist. This is particularly noticeable through the perceived child's agency in the future, when the perceived protagonist in a post is expected to shift towards this child.

> *"I would say it's more about me, my life. I post it, it is my update, my blog, I don't ask [my child's] permission. If a child in the future can turn 19 and get rights for all this information - that would be fantastic. Because this will be information about this person." [P14]*

While a particular sharing decision might still be made by a parent, the child's *ability* to express the agency seem to contribute to the perceived representation of their identity:

"As they get older, it becomes more about the kids, because they have more say at that point." [P21]

These data suggest that when a current child becomes able to express their opinion on the information's existence, previously posted information becomes more likely to be attributed to them regardless of who posted the information.

Another aspect of agency of sharing decision is the information format, which implies a choice of style and content for a post. In the context of information sharing, posting parents consider the choice of the information format as a direct reflection of their personality, and as a part of their self-representation:

"Even if I post something about what my daughter said, it's still more about me, what I find interesting. I post not just a report, but some story that touched me, that opened my eyes on something." [P15]
"I'm into photography, so when I post - it's not just information, but an interesting, artistic photo." [P14]

Correspondingly, in the context of information viewing the information format affects the perception of the social meaning of a post, and thus, a perceived protagonist:

"It is about parents, their view, what photos they choose. Three families from the same vacation would post three completely different photos of their children." [P8]
"If a parent is posting [about a child] it is always going to be about at least both of them. You as a parent choose the material, do something with it, post it. It cannot possibly be only about a child, I think." [P9]

3.5.3 Factor 3: Social connection to the recipient

The third factor affecting designation of a protagonist is the social connection to the information recipient. While communication is based on shared meanings and occurs through social coding of a message [44], social relationship becomes a shared context between a poster and a viewer of a post. When the viewer – the information recipient – is the one intended by an author, the social relationship between them and an author (sharing parent) provides the required shared meaning:

"It is me telling everyone about my child as part of my life. My friends care about him because he is mine. So, learning about him is actually learning about me." [P7]

However, when this information is received by an unintended recipient, the required shared context is missing, and the protagonist of the sharenting message is expected to change:

"It's a very fine line. Why did we decide not to post? Because right now we are posting about a child, but in 20 years it will be about them, not about us." [P4]

Similarly, in the context of information consumption, the social connection to the recipient affects the protagonist designation in "decoding" of the information in a post:

"I wouldn't think about a child in my friend's post. Because this child is not a part of this social circle. Most likely the child would not know or feel that someone has seen this photo, and what they thought." [P15]
"If I know the child, I would be more interested. I guess in this case it is more about the child, if I know them." [P23]

Furthermore, the recipient's comparative level of familiarity plays a significant role in assigning a protagonist in a message:

"If I have some kind of relationship with this child, then it is about the child. But if I only know the parent, then it's like 'oh, I see what's happening in your life'." [P14]
"I mostly assign it to the parent because I don't actually know the kids, they are a bit abstract." [P1]

3.5.4 Factor 4: Perceived agency of captured actions

The last identified factor is the perceived agency in the activity captured in the post. Specifically, in the context of information sharing, the perceived parental agency of captured actions leads to the perception of information being about themselves:

"I guess more about me, actually. For example, I went to a playground with my child. It says more about me, that I went there, than about him. He didn't take me to the playground, I took him. All these events, like, we went to a kids' center – it was my decision and my initiative." [P5]

Correspondingly, the child's perceived agency of captured actions prompts the designation of this child as a protagonist:

"A little bit of both. Actually no, I think it's more about him in my mind. 'Oh, this is something funny that [my son] did', etc. It feels more like just interesting things he does... his personality" [P1]
"More about her. Because she is only 4 month, the way she is right now it is her herself. For now I'm mainly following her - to eat, to sleep, to play. It's all her initiative." [P13]

However, the most significant role of this factor can be observed in the context of information viewing, where its variations lead to the most nuanced protagonist interpretations.

"You attribute it to whoever has the biggest identity in the post, who decided to do the action. And you usually think that parent did, so you assign it to the parent." [P1]

For example, posts about vacations or other activities, considered to be guided by parental choice, tend to be perceived as about parents, even if only a child is captured in a post:

"A photo of a child from a vacation is still about parents. Because it wasn't this child's choice to go there." [P13]
"It's not who is on the picture. Children in this case are somewhat like an attachment in the best meaning of this word. They do not decide to go for this vacation." [P8]

At the same time, information on actions which are perceived as reflecting a child's agency, personal preferences, or achievements, tend to be assigned directly to the child:

"When you are showed some personal process, like drawing, for example, I see it as a parent inviting me to get to know their child. So it's about this child." [P15]
"If a child participates in some drawing event or something like that, I consider it as about that child." [P2]
"If it's a child rehearsing a poem, or dancing - you think about the child. It's about the child's initiative." [P6]

3.6 Discussion

In our exploration of perceptions of protagonist in sharenting information, we discuss two aspects of protagonist assignment: an ambiguity in protagonist depending on perceptions of risk versus communication goals; and factors that influenced how posters and viewers of sharenting information assign protagonist. To further examine how factors and communicative aspects influence these perceptions of these information and, particularly, of protagonist assignment in sharenting posts, we combine these factors through the lenses of communication process [44] and "imagined" audiences [278].

3.6.1 Understanding Protagonist Ambiguity

Both the nature of social media, and benefits of the social disclosure that motivate posting practices [194, 40, 487, 102, 114, 246] suggest that sharenting posts have a significant communicative aspect to them. Communication as a process [44] includes a sender, who produces a message, a channel, the medium used to transmit the message, and a recipient for whom the message is intended [425]. In essence, the sender encodes the information in

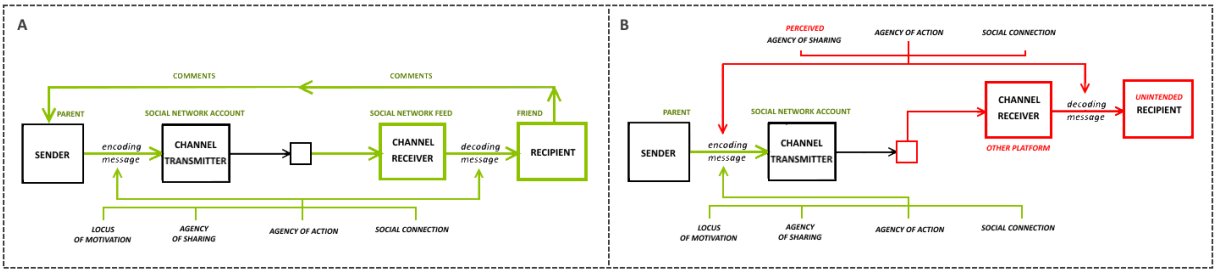

Figure 3.1: The models of sharenting in "Intended" (1A) and "Unintended" (1B) contexts

a message, sends the message through the channel, and the recipient decodes the received message. Hence, a sharenting social media post, as a communicative message, is intended – encoded – by the sharing parent for consumption – decoding – by some recipient. According to Litt [278], since on social media the actual audience is often unknown, users create a conceptualization of people with whom they communicate – the imagined audience, which determines how and what is communicated to the actual audience. The specifics of the imagined audience that a sharing parent is conceptualizing for a post determine their expectations on how, and, in particular, about *whom* this information is going to be perceived. The way how the recipient of a message (viewer of a post) decodes its social meaning depends on the understanding of the communication intention of this message [313, 281]. Thus, the expectation of a sharing parent on how the message is going to be perceived is shaped by their and viewer's shared understanding of the factors affecting decoding of a message.

Our analysis showed that depending on whether sharenting posts are considered from the perspective of current or future risk or from the perspective of communicative goals, perceptions of who a post is about (the protagonist in the post) changes. Specifically, within the intended communication context with a parent's social circle (sec.3.4.2) a protagonist tends to be assigned to a parent. However, when the imagined audience becomes less defined for a sharing parent – the present and future "unintended" audience (sec.3.4.1) – the expected shared communication context is lost, and a protagonist is more likely to be assigned to a child. Correspondingly, the undefined nature of the "unintended" audience causes parents' confusion and uncertainties on the appropriateness of what can be shared, since in this case the understanding of the "decoding" context is lacking.

To reflect these discrepancies in the designation of a protagonist and ground our further discussion, we schematically represent sharenting within "intended"(Fig.3.1.A) and "unintended"(Fig.3.1.B) contexts. For the purpose of this discussion, we build these schematic representations upon Shannon–Weaver general model of communication [424] with some

transactional elements [34].

The "intended" context of the information sharing refers to the communicative goals of a post – a parent shares a post as a social message to an expected and defined audience. The intended recipient of a message defines, for the parent, particular expectations regarding how the message is going to be perceived. The feedback, represented by comments and tokens of social appreciation (e.g. likes [227]), re-enforces the shared context, between sender (parent) and recipient, to properly decode whom the message is about:

> *"I think even comments would be not that much for the child, but for the parent." [P6]*
> *"It depends on how and where it appears. If someone sees it without comments, just a photo, then, I think, it will be more about the child. But if it appears in context, with comments, then it will be at least about both of us." [P9]*

However, when parents consider safety factors or the future perception of the information, their understanding of the imagined audience changes. The recipients of the posted information become unpredictable since the information might be directed to an unfamiliar social account or another platform, e.g. a search engine. These situations constitute the "unintended" information sharing when the understanding of "decoding" context is lacking, and the information is expected to be more tightly associated with the child and less with posting parents.

> *"I cannot predict how they are going to react. For example, I have no idea about the sociocultural atmosphere in the future - what would be normal, adequate, and what would be judged?" [P7]*

3.6.2 How Protagonist is Assigned

The four factors, presented in our results, define audience's expected and actual interpretation of which identity is represented in a post, i.e. the shared understanding of the "values" of each factor (e.g. whether locus of parent's sharing motivation is internal or external, whether agency of sharing decision is parent's or child's, etc.) form the communication context between the sharing parent and the audience. Hence, the variations between the sharenting "intended" and "unintended" contexts can be analyzed from the perspective of variations in presence of these four factors (illustrated on Fig.3.1).

Mapped to the communication process, the factor (1) of locus of sharing motivation corresponds to the parent's communicative intention – to elicit a particular response from a viewer, – and the satisfaction of this intention partly depends on the intention being recognized by the viewer [113]. According to our results, such recognition of the parent's

motivation is highly contextual, might require familiarity with the parent, and does not manifest directly in the post content. Thus, this factor is silent in the "unintended" context for protagonist interpretation, and contributes to the diversification of protagonist assignment only for the "intended" audience (Fig.3.1A).

The factor (2) of agency in a sharing decision, procedurally, corresponds to the message actually appearing online. The "value" of this factor (parent's or child's agency) is associated with the child's ability to voice their opinion on the information existence. Our results also show that, in the future, the information is expected to be attributed to the child regardless of who posted it, since the growing child is anticipated to have an ability to exercises their agency. Thus, in the "unintended" model (Fig.3.1B), this factor manifests as a social bias, often without access to the information on the actual poster.

The (3) social connection to the recipient reflects the receiving side of the process. Since a symbolic meaning of the communicated information relies on a social agreement [44, 154], social relationships secure shared context between sender and recipient. If the information recipient is from the "intended" audience, the relationship context provides the shared meaning for appropriate interpretation of the information, and, according to our results, the protagonist tends to be assigned to a parent, unless the recipient knows a child well. In the "unintended" model (Fig.3.1B), the social connection is expected to be to the child, and thus, the protagonist is assigned to the child.

Finally, the factor of the (4) agency in the captured activity is the only factor associated with the content itself, i.e. the "value" of this factor (parent's or child's agency) is defined by the information in the post rather than the interpretation of communication context; thus, it affects the diversification of protagonist assignment in both "intended" and "unintended" sharenting models (Fig.3.1).

3.6.3 Synthesis And Implications For Technical Innovation

The ambiguity and dynamics of a protagonistic identity assignment is grounded in the findings that the same information is perceived to portray a child when considered from the perspective of potential risks or the future information perception, but when considered from the intended, communication perspective, a protagonist assignment is more nuanced and gravitates towards the parent.

On one hand, this context-dependant ambiguity of a protagonist contributes to our understanding of privacy and online identity challenges associated with sharenting. On the other hand, it highlights the pressing importance of designing tools that would allow an appropriate identity segregation in information representation, that may aid users in

better contextualizing information. For example, information on a given person can be contextualized both in terms of time and social facets of one's life [250], a research direction for search result presentation that has already been explored to aid in online identity management [511]. Our results suggest that one potential solution to this problem is for social networks, communication service providers, and information retrieval systems to monitor how information from sharenting posts is re-used, sourced in posts, and retrieved in association with one's name.

Potential technical solutions for privacy issues associated with sharenting are predicated upon an understanding of how perception of such content evolves. Our results illustrate that sharenting is a prime example of posting behaviour where content perception – particularly, protagonist perception – evolves with time and within different contexts. Thus, we then presented the results of our exploration of the mechanisms of the protagonist designation in sharenting information in different contexts. We identified four overarching factors affecting the designation of a protagonist, including locus of sharing motivation, agency in sharing, social connection to the recipient of the information, and agency of the captured action. Correspondingly, we outline how and which of the identified factors affect the protagonist assignment in the "intended" and "unintended" sharenting models. These results are the first step towards the understanding of the mechanisms and structure underlying the dynamic assignment of a protagonist in sharenting posts.

From the perspective of social computing, this contribution outlines the potential directions for segregating and contextualizing each identity representation to address lasting consequences of sharenting both for parents and for children. In particular, findings on the four factors contributing to the designation of a protagonist opens exciting opportunities for developing (e.g. [27, 91]) and applying existing tools for analyzing content and metadata of posts. For example, our findings show avenues for leveraging data on social connections (e.g. [38]) between a sharing parent and the information recipient to address the differences between "intended" and "unintended" contexts. The factor of agency in information sharing, and the corresponding aspect of information format, affecting the designation of a protagonist in the "intended" sharenting model, also have a number of potential applications in context recognition. For instance, Litt [278] suggests that judging the style and content of one's communication through content analysis of one's digital traces might provide supplemental support for recognizing an author's imagined audiences. Furthermore, our findings on the factor of perceived agency in activity show opportunistic directions for the development for the recognition of social context in images [372, 373, 455, 126].

3.6.4 Limitations and Future Work

The advantages of the qualitative approach we adopt via semi-structured interviews is that the results are well-suited to developing an understandings of underlying mechanisms for both protagonist assignment and for changes in protagonist as material ages and audiences evolve. However, there are limits to qualitative data: thematic prioritization in self-reported data produced by interviews tends to relate to what participants remember, and may not reflect in-the-moment posting rationale. To reduce this limitation, we encouraged participants to open their social networking feed while discussing sharenting posts. Another example of potential limitations is that the interview data might lack granularity in the produced data, making it difficult to capture the nuanced differences in descriptions. To address this, we introduced several nuanced scenarios during the discussion of protagonist perception.

Alongside addressing methodological limitations, we also did not particularly focus our analysis on the role of affordances of social media platforms used by parents. While not necessarily directly relevant to our exploratory research questions, it would be interesting to explore the role of particular features, e.g. tagging, on the perception of a protagonist.

3.7 Conclusion

Online identity management remains a constant struggle for individuals when information appears about them online. In this project, we particularly focus on sharenting posts, a form of information created by parents, but frequently featuring their children. Through semi-structured interviews, we explore how the identity of the protagonist of these posts changes both over time and depending on perspectives taken when viewing the content. We identify four factors that influence the perception of the protagonist – locus of sharing motivation, agency in sharing, social connection to the information recipient, and agency of the captured action, – and show how these factors can potentially be leveraged to support identity management through contextual cues.

Both our work on on-line search and our work on sharenting presents instances of computer mediated communication. In the next section of this thesis, we look at person-to-person communication where this communication, while occurring directly between individuals, may be influenced or supported by information delivered by technology.

Part II

COMMUNICATION AROUND TECHNOLOGY

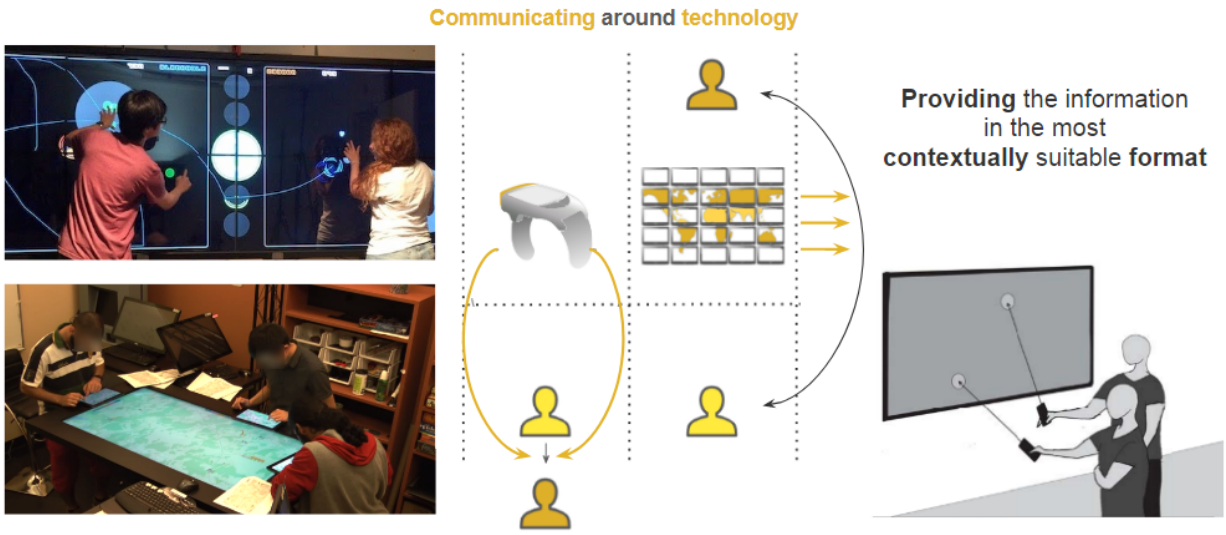

Figure 3.2: Pattern of operations adaptation in communication around the system. Examples from our work outside of this thesis (left: top [494], bottom [255]; right: [358]).

In the second part of this thesis, we explore systems that augment the user's ability to access or produce information, provoking and supporting communication *around* themselves. These systems include, for example, multi-user public displays, digital tabletops or smart boards provoking spatial interactions, or wearable monitoring technology, provoking interactions around content. Communication around the systems provides a unique set of challenges. For instance, spatial interactions imply the necessity to appropriately support users' territoriality [58, 418, 417, 26, 488], to adapt for collaborative or competitive multi-user scenarios [494, 219, 398, 460], and to support privacy considerations [79, 30, 312, 434, 128]. Furthermore, interactions around the system require supporting user's ability to manage shared physical spaces [266, 399] and appropriately interpret and respond to physical actions [185, 406, 180].

In this part, we suggest that when communication occurs **around the system**, the role of the system becomes *to provide information in the most contextually suitable format* (Fig. 3.2). We ground this suggestion in accumulated findings from two research projects(**project 3** and **project 4**).

In **research project 3** (chapter 4), we explore video-mediated participation in hybrid work meetings and analyse their prioritization and sense-making of visual attention processes. We draw our findings from two complementary studies: semi-structured interviews, exploring which aspects of visual attention contribute to participants' engagement

in meetings, and how these aspects are affected by mediated participation; and a quasi-experimental study, in which we elicited participant descriptions of attention by asking them to narrate silent video footage of work meetings to gain more granular and procedural understanding of users' sense-making of the observable information on attention (contextual meaning of pure visual signal).

Although traditionally, videoconferencing systems would be supporting communication *through* the system, specific challenges arising from hybrid meeting scenarios (when some participants are co-located and some remote) demonstrate that in hybrid meeting configuration they represent communication *around* the system. In particular, our results show that simply observable information, e.g. information on a direction of one's gaze and pose – pure signal shared with the system – is not informative on its own, but has to be augmented with a contextually interpreted function of corresponding attention processes to contribute to participant's engagement and inclusion in the meeting process. Correspondingly, our finding highlight the effects of the restrictions, that a teleconferencing system imposes on the user's ability to fully communicate contextual meaning of their actions. In essence, these restriction are based on user's inability to position their action corresponding to spatially organized objects or people. For example, addressing a person with one's gaze implies positioning the gaze direction in correspondence to spatial location of this person. Similarly, following a particular thread of conversation when multiple branches of discussion appear in a room requires user's spatially positioning themselves according to the thread of interest. Thus, we find that actions, which in non-mediated scenarios gain their contextual meanings from spacial disposition according to the other agents in the environment, loose their contextual meanings due to the limitations of traditional video-mediation.

Research project 4 (chapter 5) explores parents' use of monitoring technologies for children, including parent's information needs, motivations for acquiring this information, corresponding use-cases, and functional relationship between awareness and a sense of connectedness provided by tracking systems.

Drawing our findings from a three-phase study, which included an experience sampling study, semi-structured interviews, and an on-line data collection study, i.e. netnography, we demonstrate the tension, brought by current design of monitoring systems, between parent's resistance to pervasive monitoring and their need to monitor children to identify required actions. We show that the information of interest, which can be categorized as routine information (e.g., location, eating, sleeping), health information, school information (e.g. homework, academic performance), and social-emotional information (e.g. general mood and psychological well-being of the child), is predominantly desired when an action is required from a parent. Consequently, we identify three major use-cases for this information: *adjusting routine activities* (e.g., feeding a child, adjusting sleeping schedule),

performing emergency activities (e.g., emergency pick-up from school, finding a lost child), and *building communication and deeper personal relationship with a child*, and information format desired for these use-cases in light of parents' beliefs about privacy, trust-based relationships with their child, and a need for work-life balance.

Consolidated findings from projects 3 and 4 reveal user's tendency to resist information sharing and system support in communication when the format of information returned by the system does not properly correspond to the communication context in which this information is used. Thus, when communication occurs **around the system**, this system should be able to adapt to the communication context by providing the information in the most contextually suitable format. For instance, **project 3** shows that a teleconferencing system should adapt its operation based on augmenting pure visual signal with an interpretive meaning of attention actions, so that it can provide information to communicating endpoint in corresponding format. Then, system features aiming to support, for example, attention function of information gathering should involve timely augmentation of information streams, e.g. narrating, zooming-in, or focusing system's view or bandwidth. Supporting the communicative signaling function of attention actions requires some form of notification about one's actions delivered to other meeting participants. Finally, supporting function of following dynamic processes should include a combination of the remote participant's ability to split system views, signal their intentions, and selectively contribute to simultaneous processes. **Project 4**, in turn, shows the importance of processing and presenting the information shared with a monitoring system in light of the intended use of shared information. Specifically, on a typical day parents don't want the information available to them constantly as a stream of information on what is happening right now. Instead parents prefer summary information, a recounting of what has happened during the day apart. Parents then would use this information to adjust their behavior when with a child later that day, or as a basis for shared activities.

The following two chapters present a detailed discussion of **project 3** and **project 3**.

Chapter 4

Project 3: Functional Classification of Attention in Video Meetings

In television and other video-based entertainment media, a primary goal is to appropriately capture and guide the attention of the viewer. Unfortunately, in video conferencing, it is challenging for users to attend for long periods of time, primarily because our video conferencing systems typically show us a single view of remote participants, without taking into account objects of interest and optimizing views around those objects of interest. Even when features such as screen sharing are used, the user must explicitly manipulate views, and automation is rare. In this chapter, we look at whether viewers of meeting videos can identify primary artifacts of interest as a first step in understanding how to leverage attention identification to better guide attention during video conferencing sessions.

4.1 Introduction

Video communication is well known for the trouble people have with getting and paying attention [190, 369]. This is especially the case for hybrid work meetings in which participants are distributed unevenly over endpoints. Given that many mechanisms of situational awareness [199, 190] and social presence [210, 48] rely on visual cues, it seems especially galling that the value of the video channel ends up lost in translation [141, 174, 200, 430]. Attention's dual nature as both interactional and cognitive [221, 63] seems especially vulnerable to video communication's decades-long struggle with asymmetrical situational awareness [501, 153, 196].

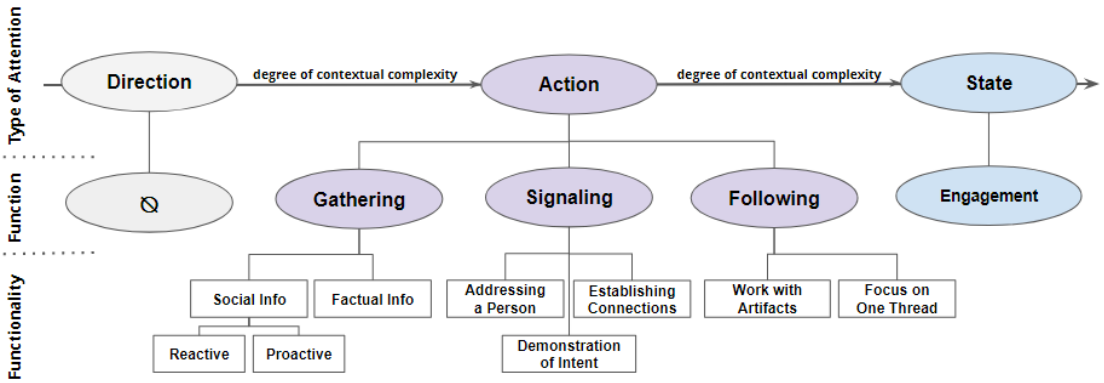

Figure 4.1: Classification of functional attention

Research on attention has ranged from gaze [181], through joint frame of reference [172] and situated awareness [180, 172], to broader concepts of engagement [47, 404]. There is a correspondingly extensive body of targeted solutions for providing situational cues for attention [480]. For example, systems have aimed to mitigate limitations of remote participation through tracking focus of attention [446] and eye-contact [144], selectively focusing bandwidth on participants' faces [515], supporting spatial arrangement similar to face-to-face meetings [527, 421], enabling remote users to point at objects within environments [345, 346], providing feedback on the conversation to participants [407], or immersive meeting environments [348, 165, 527, 245]. In sum, research on attention is very fractured [212]. Thus, there is a need for systematic understanding of what it actually means to support remote user's attention.

In this project, we analyse participants' prioritization and sense-making of visual attention during video meetings at work. We focus on visual attention since video is proposed to support the many mechanisms of situational awareness and social presence rely on visual cues, because visual information is considered to be particularly important as a focus of attentional resources [132]. Our findings are drawn from two complementary studies. We first conducted semi-structured interviews to explore which aspects of visual attention contribute to participants' engagement in meetings, and how these aspects are affected by remote participation. To triangulate and further probe themes emerging from the interviews, we conducted a second, quasi-experimental, study in which we elicited participant descriptions of attention by asking them to narrate silent video footage of work meetings. These results provided us with more granular and procedural understanding of users' sense-making of the observable information on attention.

We synthesize our findings into a three-level functional classification of visual attention

in meetings (see fig.4.1).

- The top level represents three *categories of attention* that a video meeting feature might support. First, *Attention as Direction* – purely observable attention that constitutes a process, but not a recognized purpose (e.g. Alice is looking at Bob). Second, *Attention as Action*, which reflects attention processes with a recognized specific purpose (e.g. Alice is addressing Bob). Third, *Attention as State* – the overall sense of engagement, which reflects only purpose, but no particular process (e.g. Alice is attentive).
- The middle level represents *functions of attention*, i.e. recognizable purposes: There are three functions of *Attention as Action*: Gathering Information ('look to see'), Communicative Signaling ('look to be seen'), and Following Dynamic Processes ('look to be a part of a process'). A general Sense of Engagement ('be mentally there') is a function of *Attention as State*.
- Finally, in correspondence to each of three functions of visual Attention as Action, we elicited specific attention-related meeting activities, which we refer to as *functionalities*. Organized according to the corresponding attention functions, these functionalities form the bottom level of the classification.

Essentially, our model takes the amorphous concept of attention and develops a multi-level representation, focusing on what it means to **support** users' attention during a meeting. Attention actions are characterized by their procedural aspect (those amenable to technological augmentation) and functional aspect (meeting activities) (fig.4.1). The intent of our model is provide a language that allows designers to be more deliberate when deciding and articulating what type of attention they wish to foster through their tools.

For example, consider the 'active speaker view' common in many video meeting systems (e.g. [117]). This feature uses audio to trigger enlarging or highlighting whoever is speaking at any given time and automatically switches the focus when a new person starts talking. According to our classification, this feature supports *Attention as Action* with the Information Gathering function. Another example is Companion Experiences for Microsoft Teams [389] which allows users to pair their computer and phone during a group call, for example, to present live mobile video or a photo while simultaneously maintaining the call on the computer. The Companion feature supports *Attention as Action* with the function of Process Following because it allows the presenting user to focus on an artifact in the meeting environment without losing social and situational awareness of the meeting in general. Finally, eye-tracking systems that correct remote users' perceived gaze direction [348], support *Attention as Direction*.

Our results reveal that the ability to perform purposeful attention actions significantly contributes to the participants' overall sense of engagement. We suggest that *Attention as Action* yields the most promising initial direction for feature development. Arguably, though, the ultimate goal of features supporting user attention is to foster *Attention as State* – keeping remote participants generally engaged and mentally 'checked in'.

In the sections to follow, we first review previous research on attention in computer-mediated communication contexts, particularly highlighting visual attention issues. We then present our methodological approach and describe the design and results for each of our studies. Finally, we synthesize our findings as a functional classification of visual attention, discuss it in context of other theoretical conceptualizations of attention and situational awareness, and outline the implications for feature design.

4.2 Previous Research

Attention is above all a progressive process of selection that lies as much in our sense-making abilities as it does in our physical abilities. Research on computer-mediated communication have explored this notion of attention from a variety of standpoints.

Telecommunication studies of the 1970s considered the disadvantaged position of remote users to be a function of bandwidth of a medium's channels. Building on the concept of immediacy as a set of behaviors which increase involvement [309], Short et al.[428] introduced the concept of social presence to the context of telecommunication. They argued that social presence is a quality of medium itself and is enhanced by immediacy conveyed both verbally and non-verbally. From the 1980s onwards, explorations of social presence in video meetings argued that social presence depended not just on the medium itself (text, audio, or video) [337], but on interactions enabled and supported by this medium [110, 421, 479, 112]. Picard [357] suggested a focus on an "affective channel capacity" - how much emotional information a channel can carry as compared to overall information. In the early 2000s research emphasis changed from a bandwidth orientation to a broader sense of social presence. For example, the Networked Minds Social Presence framework [48, 47] suggests that social presence is composed of three levels. A base, perceptual level is characterized by a sense of co-presence with the embodied other including observations of the other's identity, intentions, and attention. The middle subjective level represents the psycho-behavioral accessibility of the other, sense of connection to their intention, attention, and affective state. The top, intersubjective level reflects a sense of behavioral engagement when actions are perceived interdependent with actions of the other.

Parallel research over similar periods emphasized the importance of shared environmental awareness. In video communication research, this meant considering the need for more views of remote endpoints than the traditional face-to-face views, even as these extra views led to problems with establishing a joint frame of reference [172, 200, 288]. These problems stem as much from cognitive limitations as they do from interface issues [214]. The cognitive argument here is that while the environment is an extremely rich source of information, not all this information is equally important, and selectively attending to the information filtered as relevant helps communicators to establish shared knowledge [430] and avoid cognitive overload [200]. One of the mechanisms supporting and informing such prioritization involves directing more cognitive resources to targets of shared attention [430]. Shared or joint attention is a result of an ability and motivation to follow the direction of gaze or pointing actions [316] and develops based on social awareness as signaled by physical proximity, head and body orientation, as well as dynamic gaze [362].

Gaze itself has been extensively studied as an attention mechanism [480, 212]. Research in social cognitive science shows that observers of complex scenes select the people's eyes as it allows observers to derive social information, more than bodies, background, or foreground objects [49, 50]. Further, people look at the eyes of others more frequently if the scene is highly social and includes several people engaged in joint activity [49, 50]. See Frischen et al. [162] for an overview of past research on the perception of gaze behavior and its effect on the observer. Eyes both collect information and communicate one's state [21] and attentional focus [162]. The dual nature of gaze has been adopted by modern cognitive science to explore naturalistic social attention [390]. For instance, in collaborative processes, the coupling of gaze patterns between interlocutors has been found to be causally related to the knowledge people share before a conversation and the information they later recall [387]. Gaze thus supports establishing common ground [20, 100], for example by disambiguating references [193] and helping to predict next task actions [520].

Video-mediated communication researchers have explored various non-verbal cues [446], with many focusing on detecting gaze [348] and eye-contact [144], or simulating gaze [345, 346, 512]. Gaze coordination has also been extensively explored in human-robot interaction [67, 311, 391, 404]. For example, people understand robot speech better when a robot's real time gaze behavior is similar to that of humans [441], and when a robot is interacting with multiple people, the visual focus of attention affects addressee recognition [427].

While gaze is tightly related to general and social attention processes, social cognitive studies shows the specifics of this relationship are not necessarily linear. Gaze might mean attention but attention does not always mean gaze. For instance, analyzing participant's viewing behavior during social interactions, Freeth and colleagues [161] demonstrated that people look at the partner's face significantly less when answering a question compared

to when they are listening to the question being asked, which supported earlier research showing that averting one's gaze from other people can help one to think more effectively by reducing visual processing demands and cognitive load [127, 177]. Thus, while gaze indicates focus of visual attention it does not necessarily reflect focus of cognitive attention. Further, for gaze to play a signaling function [21] and invite joint attention in communication [405, 429] an interlocutor needs to recognize not only the focus of directed attention, but also the environmental context for such attention [180].

The range of different approaches to attention and technologies created to overcome attention problems suggests that designing corresponding features requires a better granular baseline for understanding what aspects and processes of attention should be augmented by technologies. Our research aims to identify the fundamental types of attention in meetings, accessible for technological support, and how they might be recognized during meetings. Thus, we explore the functional aspects of attention, i.e. how and why people pay and perceive attention.

4.3 Method

To probe participants' prioritization and sense-making of attention processes, we conducted two studies. First, we interviewed participants about attention in work meetings. We then quasi-experimentally elicited descriptions of videos of work meetings to triangulate themes emerging from the interviews.

We took a qualitative approach in both studies because our goal was to explore the way that participants understood attention and prioritized contextual cues. Interview-based research on video-mediated communication tends to report results that are well-suited to developing thematic glosses and evaluations. However, while people are generally good at providing narratives, they tend to be less successful at remembering the granular cues from which they weave these narratives. Additionally, thematic prioritization in self-reported information tends to relate to what is remembered. On the other hand, experimental procedures develop more replicable results that lend themselves to comparison and contrast, as well as tracing flows of action, but lack contextual richness and require deciding *a priori* on a subset of observable measures. Hence, for this work, we started with semi-structured interviews to gather open-ended accounts of meaningful attention, followed by a quasi-experimental study to act as a focus-check and expansion of those accounts. This data triangulation allowed us to cross-validate our observations between the two studies.

4.3.1 Study I Design

Through a series of semi-structured interviews [496], we explored users' experiences of participating in work meetings of various configurations, with a particular focus on attention processes and engagement. In total, we interviewed twelve participants (age 21-50, 5f,7m). All had a tertiary degree in diverse areas of knowledge, including engineering, science, and arts[1]. All participants had extensive experience in co-located, remote, and hybrid meetings (at least once a month for at least over a year), and each participant was a full-time employee at the time of the study.

The interviews were conducted one-on-one, in-person, and informed consent was obtained prior to the interview. Each audio recorded interview lasted approximately 40 minutes. We started by asking participants about their preferences for meeting configurations and what motivated those preferences. Through these discussions we also explored their practices and expectations for joining meetings in person and remotely, with video or audio-only channels available. We then concentrated on remote meeting participation experiences, asking participants to describe the strategies they use to understand a meeting environment and social dynamics during a meeting, strategies to recognize and support their own and other participants' attention and engagement, and obstacles related to attention dynamics that remote participants might experience.

After transcribing audio recordings of each interview, we performed incremental data analysis across participants using open and axial coding [107]. Each phase of coding was initially performed by the primary author, and then codes were discussed and refactored in consultation with the research team.

4.3.2 Study II Design

To validate and extend the granularity of the first study results, we conducted a second, quasi-experimental study, asking participants to narrate muted videos of work meetings, focusing on attention processes they observe. We recruited a new set of fifteen participants (21-36 y.o, 8f, 7m), intentionally excluding those who participated in the first study. We wanted to prevent a potential primacy bias and to support the external validity of the results by extending and diversifying the participants pool. While the goal of the first study was to generally explore participants' experiences regarding attention and engagement

[1]Note that individuals who engage in hybrid work meetings are often in professions that require tertiary degrees.

during work meetings, the second study focused on further understanding the prioritization and sense-making of purely visually observable cues of attention.

Participants were presented with six short video records (2 minutes) of real work meetings. Meeting scenarios included 4 to 6 meeting participants, either all co-located, or in a hybrid meeting configuration. The video clips reflected different combinations of the following meeting processes: joining and leaving a meeting environment, setting up a meeting environment, talking around a table, taking notes, using a whiteboard, using a projector screen, using smart devices. The videos were presented with muted audio, and participants were asked to narrate the visual information from the video to an experimenter sitting next to them. We chose the muted condition for two reasons. First, we wanted to eliminate potential biases on information interpretations based on the audio context. Second, we wanted participants to focus on the purely observable information and expected that the lack of audio context would ensure participants' focus on visual cues. At the beginning of the experiment participants practiced using an additional training video clip of a work meeting. After the experiment, we conducted a debrief interview asking participants about their experience and strategies for information descriptions and prioritization.

This second study was quasi-experimental in that we controlled the stimulus and required participants to follow a particular protocol, but we did not go so far as to develop and test hypotheses. Our analysis focused on participants' verbal descriptions of attention processes. From it, we developed a set of 'attention identifiers' – distinct meaningful elements of a description that signified an act or process of attention. Attention identifiers were most commonly denoted with a verb, examples of which included "to look", "to watch", "to listen", "to pay attention", "to direct attention", "to be attentive to", "to be engaged", "to have a conversation with", etc. In total, we identified and semantically analyzed 225 responses of observable attention from 90 descriptions of 6 videos.

4.4 Study I Results

In this section, we present our interview findings (fig.4.2). We explore what functional aspects of attention contributed to people's sense of meeting engagement, and categorize them as Gathering Social Information, Following Dynamic Processes, and Communicative Signaling. For each function we describe the elicited activities in the meeting process.

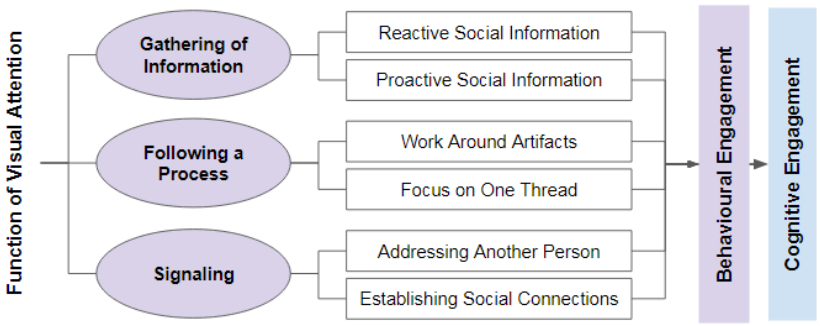

Figure 4.2: Results of the Interview Study (Study I)

4.4.1 Functions of Visual Attention in Engagement

Exploring the role of visual attention in users' sense of engagement, we first examined what information participants generally perceive as relevant and pay attention to during meetings. We then cross-analysed it with the information that participants perceive as more difficult to acquire in different meeting configurations. We found that, despite having a video channel, remote participation is perceived to affect user's access to three functions of visual attention (fig.4.2).

Gathering Social Information

The first function of visual attention affected by remote participation is Gathering Social Information, which refers to the ability to support social situational awareness. Even with a video channel, remote participation is perceived to constrain the user's ability to collect social information required for suitable behaviour adjustments.

In our analysis, we distinguish Gathering Social Information as reactive or proactive. Gathering *Reactive* Social Information indicates responses to one's actions and used to assess the social environment and required adjustments. Gathering *Proactive* Social Information indicates the social disposition in the environment and required to initiate one's action. Correspondingly, two themes emerged in participants' concerns regarding the accessibility of social information in remote participation: gathering of information on emotional reactions of others (reactive), and assessing social-behavioral cues to manage their own participation (proactive).

Limitations on Gathering *Reactive* Social Information affect the user's ability to assess the reception of their actions, and, thus, the ability to appropriately adjust their social behaviour:

"It's about being able to see the other person, their facial expressions, body language very clearly, which you can't always do even on the video. It is easier [in person] to get a feel for how the other person is reacting." [P12]
"If I could see everyone, which we can't usually do, but let's say we could, so I can present and see how they are reacting, I don't know if I would need to be there in person." [P9]

In addition to facial expressions and body language, this also includes more complex social-emotional cues, e.g. other participants' engagement and attention to one's actions.

"I'm looking for people nodding or smiling... I might rely on it subconsciously to see if people are engaged or not, to see whether people look interested or excited." [P4]
"During a Skype meeting I feel like it's difficult to say whether someone is paying attention." [P2]

Limitations on Gathering *Proactive* Social Information affect the remote participant's ability to appropriately initiate their action according to the current social environment. This, first, includes the information required to inform appropriate speech patterns – social cues on when is it appropriate to start talking:

"When remote, it's an additional challenge - you don't see cues that will let you say 'I'm gonna speak now'." [P9]
"In person you have all these non-verbal cues like you can see when someone is about to say something." [P10]

It also refers to gathering information on who is in the room, their social roles, and social dynamics between participants:

"It is also hard to figure out who is important, what's happening in the room." [P3]
"The problem is that you don't know who is looking at you. When you are there, you can move your head, see what's going on. On camera you have no idea." [P9]

Following Dynamic Processes

The second function of visual attention, Following Dynamic Processes, refers to the user's ability to dynamically direct their attention to follow the meeting's work. Two themes emerged in relation to this function: participation in work processes around environmental artifacts, and navigation within multiple conversation threads.

Following Dynamic Processes around environmental artifacts mainly included well established difficulties of working with whiteboards and projected presentations:

"I was joining remotely, others were in person. I was clearly sensing disadvantage. Because people were working on a white board and it was difficult to see." [P6]
"It's hard if there are slides or other materials... maybe I would like some indication of what's going on." [P1]

Another related challenge is the ability to follow one of multiple conversation threads, especially in hybrid settings:

"If I am remote and multiple threads of discussion develop in the meeting room, then it's like 'how do I get streamlined into one thread of discussion'." [P7]
"If a big meeting splits into several conversations, I just switch it off all together, because I can't join one of those conversations. But if I am in the room I can turn to my neighbour, have a subgroup." [P11]

Communicative Signaling

Lastly, remote participation affected the Communicative Signaling function of visual attention. While people are gathering and following, they also signal their reactions and intentions as part of communication, i.e. an act of visibly paying attention is itself social message to other meeting participants. The Communicative Signaling function contributes to a group communication dynamics. For instance, a constrained ability to use Communicative Signaling results in difficulties of directly addressing others:

*"Even just looking, talking to that person directly makes a huge difference. And it would be nice to know that people are speaking to *me*." [P2]*
"If somebody addresses me, in person they turn, face me. In a remote meeting they would have to turn to the camera, or say my name. All these additional decorations have to happen to ensure that I am included adequately." [P4]

Constraints on the Communicative Signaling function are also reflected in the user's reduced ability to manage their social interactions and, correspondingly, to appropriately participate in the meeting dynamics:

"In person, I make sure that I'm not just talking to one person or one direction in the room. You need to switch focus, talk to everyone. That's not something I can control remotely – I'm only looking at whatever I can see." [P3]
"When in person, you can look around the room, make contact with people who aren't talking. You know, make connections with other people. You can't do it remotely, you are only looking at one thing." [P5]

Similar to the gathering and following functions, the ability to use the Communicative Signaling function of attention contributes to the overall sense of engagement, both for a remote participant themselves, and for the other participants' perceptions:

> *"If a person is not trying to engage the crowd, then I lower my own engagement. I just kinda put it on the background. It's hard to engage with someone who isn't making an eye contact at least with someone." [P3]*
> *"On Skype, you have video, you can notice when somebody is not paying attention... and you feel like people are watching you, so you can't stop paying attention." [P9]*

4.4.2 Section Summary

To summarize, we explored the role of visual attention in a remote participant's engagement and found that people describe remote participation as affecting distinct functions: Gathering Social Information, Following Dynamic Processes, and Communicative Signalling. Each of these functions contributes both to people's general sense of engagement and is also associated with corresponding functionalities: assessing reactions of others and initiating one's own participation, working around artifacts, focusing on one of multiple discussion threads, addressing specific participants, and establishing social micro-connections.

4.5 Study II Results

To further explore the granular prioritization and sense-making of attention processes, we conducted a second study in which we asked participants to narrate the attention processes they observed in silent videos of work meetings. This resultant attention descriptions were, of course, quite heterogeneous (fig.4.3), but we found a useful thematic distinction between purely descriptive and functionally interpreted attention. The results of second study validate our earlier findings on three functions of visual attention, and expand the set of related functionalities.

We first, present our analysis of 'attention identifiers' – distinct descriptive elements signifying an act or process of attention. We then deepen into the participants' contextual interpretations of attention processes reflected in these identifiers, and describe the elicited types of interpreted attention, followed by the analysis of their interpreted functions.

4.5.1 Types of Attention Descriptions

We began by analyzing the differences in the attention identifiers used by participants to denote attention-related processes. Based on this analysis, we categorized attention identifiers as either descriptive or indicative (fig. 4.3). The descriptive identifiers were 'objectively' translations of visual information into a verbal form ("looking at an object"), representing narrations of purely observable attention. The indicative identifiers were triggered by the same visual information, but were formulations of the observer's interpretation of the social relevance of the visual cues.

Descriptive Attention Identifiers

Descriptive attention identifiers were 'objective' narratives of visual information, simply indicating the direction of gaze or body orientation (e.g. *'now both guys are looking into the phone'* [V6P10]). They were 'objective' in the sense that they did not contain any information on the purpose of the corresponding attention processes, but only registered their occurrence.

For example, in the description: *"The guy with the long hair is looking at the guy with the Coke"* [V3P10], the descriptive attention identifier "looking at" purely denotes that the gaze of "the guy with the long hair" is directed towards "the guy with the Coke". For further illustration consider the following examples with descriptive attention identifiers: *"Man number 2 is* ***looking at*** *the whiteboard"* [V4P6]; *"Person 2 is laughing and their* ***attention was just directed*** *to the screen"* [V1P4]. The type of observable attention signified using descriptive identifiers is purely registrable, thus it relates the function of Attention as Direction.

Indicative Attention Identifiers

In contrast to the 'objective' nature of descriptive identifiers, indicative attention identifiers include narrators' contextual interpretations of observable information (e.g. *"people still seem to be paying quite close attention to the speaker"* [V3P7]). Thus, the indicative identifiers 'indicate' the functional aspect of the corresponding attention process, according to the narrator's interpretation. For example, the description *"he is getting involved in this conversation with the person on Skype"* [V6P1] narrates the scene of a man looking at the screen displaying another person. In this description, the indicative attention identifier *"getting involved"* is formed based on the visual information which is contextually interpreted as engaging into a Skype conversation. While indicative identifiers are

still triggered by atomic visual stimuli (recall that participants were presented with muted videos), the corresponding attention descriptions are produced as declarative as opposed to purely descriptive: *"They are **listening** to the woman... **listening intently**."* [V2P5]; *"The women is **responding** to what the guy was saying"* [V5P14].

Participants would also commonly augment descriptions of attention processes by interpreting related social dynamics: *"The person who is videoconferencing is speaking because everybody is looking at the screen to listen to this person"* [V4P8]; *"She must be explaining something because everyone is looking at her and nodding as well"* [V2P9].

This tendency to augment descriptions with contextual interpretations was further reflected in data from debrief interviews:

> *"I interpret the gesture as having some meaning in a sort of conversational dialog, back and forth. It was both the captured motion and then the meaning of the motion that I was interpreting." [DebriefP3]*
> *"I also tried sometimes to add additional features like 'they are doing some actions' but not all the actions, I tried to find the leading person every time, who is leading, who is speaking, who is drawing - because I find that that's how the pipeline of the meeting is happening." [DebriefP12]*

Participants omitted the functional augmentations of attention descriptions if the communication context was missing or unclear. In this case, they would resort to the descriptive identifiers to merely capture the direction of attention.

> *"That could be tricky because sometimes it's not clear whether there is something currently being spoken about which warns them to look at something." [DebriefP6]*

The 'interpretive' nature of indicative identifiers constitutes the functional aspects of the observed attention processes – their perceived contextual meaning.

4.5.2 Types of Functionally Interpreted Attention

Indicative attention descriptions showed that corresponding contextual interpretations included either descriptions of attention as cognitive state (e.g. *"they seem to be quite engaged"* [V5P12]), or descriptions of attention as a purposeful action (e.g. *"the blond woman is listening to the conversation he is having with the woman with dark hair"* [V5P10]).

While descriptions of attention as a cognitive state were triggered by visual information, they typically did not include information on related behavioural displays. For example, in the description *"They seem to be quite engaged"* [V5P12], the indicative attention identifier

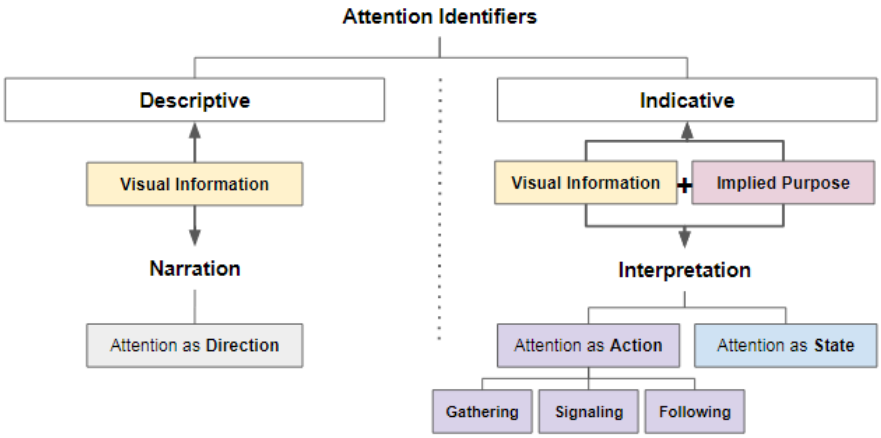

Figure 4.3: Results of the Narration Experiment Study (Study II)

"to be engaged" describes the overall state rather than a specific behaviour. Moreover, these descriptions do not specify a direction or a target of attention: *"Otherwise people seem to be pretty attentive"* [V3P7]; *"The man is not paying attention anymore"* [V5P10].

In contrast to the descriptions of attention as a cognitive state, descriptions of Attention as Action include a particular direction and a specified target of an attention process. For example, in the description *"person 4 is talking again to persons 1 and 2"* [V5P4], the indicative attention identifier "talking to" implies a direction of the attention act toward a target – "persons 1 and 2". Similarly, the following examples of Attention as Action include a target, denoted in both cases by a pronoun "him": *"The guy in glasses is talking and everyone is listening to him"* [V3P10]; *"Now person 2 on the right is talking and the other two are paying attention to him"* [V1P1].

Furthermore, descriptions of Attention as Action were formed based on an interpreted purpose of observed actions. For example, *"listening"* is a contextual interpretation of atomic visual information, such as *"looking in a direction"* of someone who is speaking. For more illustration, consider the following descriptions: *"Really it's two people talking to each other and other two are just kind of watching them"* [V2P7]; *"The blond woman is just listening to the conversation he is having with the woman with dark hair"* [V5P10].

Informed by the first study results, we analyzed these inferred intentions in terms of the three functions of our model: actions for Gathering Information, actions intended for Communication Signalling, and actions for Following Dynamic Processes.

Gathering Information

The attention actions for Gathering Social Information can be described as 'look to see': e.g. *"Man 1 is checking the screen of the smart device used by man 2"* [V6P6]; *"The person who is standing kind of looks at what he'd drawn"* [V4P7]. While data from the first study revealed the theme of Gathering Social Information (sec.4.4.1), the quasi-experimental data extends those results by demonstrating that all types of environmental information are noticed and appears in description of the attention processes. e.g. *"They are talking about something... yeah... the two guys at the back don't, they are just listening"* [V2P2].

Communicative Signalling

If actions for Gathering Information can be described as 'look to see', then actions for Communicative Signalling are the converse: 'look to be seen'. These actions are performed to be acknowledged by other meeting participants, e.g. *"The woman is talking. She is addressing both men."* [V1P6]. Communicative Signalling typically refers to either displaying attention to a process (e.g. *"person 2 on the right is talking and the other two are paying attention to him"* [V1P1]), or to directing speech to specific meeting participant (e.g. *"Man 3 is addressing everyone at the table, making eye-contact with everyone"* [V3P6]).

Our interview results identified two main functionalities for Communicative Signalling: establishing social micro-connections with other meeting participants, and directly addressing specific people (sec.4.4.1). Again, While validating these results, the quasi-experimental data also expands them by revealing a functionality of demonstrating intent, when a meeting participant displays their engagement in the process, e.g. *"he doesn't seem to be acknowledging the conversation, he is kinda writing something down in his notebook"* [V5P4].

Following Dynamic Processes

The third type of inferred purposes of attention actions is Following Dynamic Processes, describable as 'look to be a part of [a process]'. These actions combine the Gathering Information and Communicative Signaling purposes with an addition of a dynamic target of the action (e.g. adding information on a whiteboard). For example, *"They are adding drawings to the board, looking at it intently. And discussing the best way to do this"* [V4P11]; *"Everyone is very interested in what this person is about to do on the board"* [V6P3]; *"he is talking to the guy who is drawing, instructing him on how to draw stuff or something like that"* [P2V6]. Additionally, these actions require a person to be able

to dynamically switch between a complex combination of targets, including social and environmental targets, e.g. *"they sometimes refer to the screen and say something, discuss something that's displayed on the screen"* [V1P14].

Thus, the quasi-experimental data supports one of the two themes on the ability to direct remote participant's attention from study I: obstacles to participating in work processes involving environmental artifacts (sec.4.4.1). The theme of navigating within multiple conversations did not appear in the quasi-experimental narrations, probably due to silent nature of the video stimuli.

4.5.3 Inter-dependency of Functional Types of Attention

Finally, we explored the inter-dependencies between descriptions of one's attention as cognitive state and attention as a purposeful actions. Validating the first study's results, the second study debrief interviews also suggested that cognitive engagement is significantly influenced by one's ability to perform specific attention actions. This, for example, includes the ability to gather information on other participants' attention-related actions, and to noticeably (behaviourally) participate in the meeting process:

> *"what was important to the members of the meeting – what they would be looking at, carrying about. Rather than details on what's someone is not doing or what is not being looked at - it's not relevant. The contribution to the meeting process is critical, the ones who don't speak or appear to be paying attention tend to be kind of ignored." [DebriefP6]*

The ability to noticeably participate in the meeting process also affects the social inclusion of the participant:

> *"If someone doesn't say or do anything, you can just forget them." [DebriefP2]*
> *"Maybe someone walked into the room but they were not talking and were not noticed as a new player by the people who are already in the meeting – like someone else is already having a conversation and a new person came in and they didn't immediately acknowledged that person coming in - then it wasn't as important, as if that person came in and started talking." [DebriefP4]*

4.5.4 Section Summary

The analysis of attention descriptions allowed us to discriminate between three types of noticeable attention in the user's sense-making: Attention as Direction, Attention as Action, and Attention as State. We categorized the interpreted intentions associated with

attention actions as Gathering Information, Communicative Signaling, and Following Dynamic Processes, which validated our earlier results of three functions of visual attention. Finally, the results from both our studies suggest that the ability to perform specific attention actions and, thus, the accessibility of corresponding attention functions contribute to the participant's overall sense of engagement in a meeting.

To highlight these commonalities and to operationalize them for feature design in video meetings, in the following section we synthesize our findings as a three-level functional classification of observable attention.

4.6 Functional Classification of Attention

Based on cross-validation of the results from the interview study and the analysis of the narration experiment data (fig.4.2 and 4.3), we constructed a three-level classification of functional types of observable attention in work meetings (fig.4.1). The top level represents functional types of attention, organized according to the degree of the contextual complexity required to recognize and interpret each type of attention. The second level represents functions of attention types, affected in video meetings. The bottom level organizes the functionalities of attention actions – tasks and activities within a work meeting, – corresponding to the second-level functions of attention.

4.6.1 Level 1: Contextual Complexity Spectrum

First, the classification includes three types of attention, accessible for video meeting feature support. These types of attention are organized on a spectrum left to right, from the least to the most contextually complex for recognition.

Attention as Direction represent the contextually simplest type of noticeable attention processes without a recognized purpose. In fact, both our studies demonstrate that Attention as Direction cannot be fully considered as a functional type of attention. Instead, it encompasses purely observable attention processes, which, in turn, can be used to infer social meanings for either Attention as Action, or Attention as State.

For instance, the interview participants acknowledged that they considered the visible information as Attention as Direction, but also articulated that these visual cues are not fully informative for assessment of one's attention actions and engagement:

"If it feels like I am looking at you and I'm like 'no, I'm not, I just got an email on the screen. So you only see if they are looking at the screen or not." [P2]
"because camera is over here and they are looking at the screen, so I can't tell whether they are actually looking at me or at something completely different." [P9]

Similarly, the second study shows that descriptions of Attention as Direction might, in fact, be incomplete descriptions of Attention as Action, when contextual information required for interpretations is missing or unclear (sec.4.5.1).

Attention as Action refers to the attention processes with specified direction and target, as well as recognizable functional meaning (purpose). For instance, in our interview data, Attention as Action would be described as observable but interpreted actions: *"people are looking at you, listening to you"* [P1]; *"you can know by looking at someone who is making eye contact when talking"* [P3]; *"them reading their phone"* [P9]. We positioned Attention as Action in the middle of the contextual complexity spectrum, between Attention as Direction and Attention as State. While Attention as Action augments Attention as Direction with the functional meaning of the processes, it is less contextually complex than Attention as State. First, structurally, descriptions of Attention as Action include a direction and a target of an attention process, which makes them easier to recognize in the environment, and second, because it causally affects Attention as State – the ability to perform attention actions contribute to the overall sense of engagement.

Lastly, the most contextually complex, is *Attention as State*, which reflects the overall sense of one's engagement in the meeting process. In other words, when we casually talk about paying attention to the meeting and being engaged, we generally refer to the Attention as State. For example, in our interview data, Attention as State would be described as general engagement or *"directing all my attention [to the meeting]"* [P11], while the lack of Attention as State would be described as *"I don't have 100% of my attention available"* [P7]; *'I check out mentally'* [P10], or *"put [the meeting] on the background"* [P3].

Our classification positions the functional type of Attention as State on the far right end of the contextual complexity spectrum for several reasons. First, Attention as State seems to be directly affected by the accessibility of Attention as Action. Second, the structural analysis of corresponding attention descriptions shows that descriptions of Attention as State do not include a specified direction or target of an attention process (e.g. being *"pretty attentive"* [V3P7]). Thus, Attention as State reflects only purpose, but no particular process associated with it, which makes it the most complex to recognize and support in the environment.

4.6.2 Levels 2 and 3: Functions and Functionalities

Since Attention as Direction refers to purely observable attention that constitutes a process, but not a recognized purpose, the corresponding second and third levels of the classification are empty (fig.4.1, left). Attention as State only has a general purpose of being engaged (fig.4.1, right). Finally, Attention as Action (fig.4.1, middle) augments Attention as Direction by having both a process and a purpose. The corresponding second level of the classification reflect three identified functions of attention: Gathering Information, Communicative Signaling, and Following Dynamic Processes. Each of these functions corresponds to a set of functionalities within a work meeting – specific tasks and activities performed through the attention actions, which form the third level of the classification.

Gathering Function

First functional type of attention actions includes the actions directed to the passive collection of information from the environment. In the interview results, information gathering mostly described collecting information for social situation awareness: either reactive or proactive social information. The experiment data, however, allowed us to expand this category by including the collection of 'factual' information as well.

The functionality of passively Gathering Reactive Social Information refers to the ability to assess the reception of one's communication action by the other meeting participants, e.g. *"If you are doing a presentation for a client, you want to know how they are receiving things, when they are losing attention, when they do take notes."* [P5]. The functionality of passively Gathering Proactive Social Information reflects the ability to assess the social situation to appropriately initiate one's actions, e.g. *"Cues as to when people look like they can be interrupted, it's a bit harder when you are doing it remotely."* [P4]. Finally, the functionality of Gathering 'Factual' Environmental Information refers to one's ability to passively collect the non-social, content-related information during the meeting process (e.g. *"checking the screen"* [V6P6], *"just listening"* [V2P2]).

Communicative Signaling

The second functional type of attention actions includes actions performed to be acknowledged by others. Both studies' results identified two main functionalities within the signaling function of visual attention: establishing social micro-connections with other meeting participants, and directly addressing specific people. The analysis of the quasi-experimental

data allowed to add the functionality of demonstrating one's engagement in the process, which was further supported by revisiting the interview data.

The functionality of establishing social micro-connections allows meeting participants to *"make contact with other people who aren't talking"* [P5] as well as to gain attention of others, e.g. *"In person we rely on quite a lot eye contact and people on calls, they tend to do a lot to make sure that they grab the attention of the one person that they specifically need."* [P2]. The functionality of addressing a person reflects one's ability to direct their message to a specific subset of meeting participants: *"It would be great if you could direct your conversation towards a certain group of people... like 'I want to talk to A and B'. But yeah... that's not really possible."* [P11]. Demonstrating one's engagement refers to one's ability to purposefully display their involvement in the process, e.g. *"you can know if they are paying attention by looking at who is making eye contact when talking"* [P3].

Following Dynamic Processes

Finally, attention actions with the purpose of Following Dynamic Processes combine information gathering and signaling functions with addition of the dynamic nature of the target. This means that Following Dynamic Processes is directed to the dynamic changes within the immediate environment, which consequently allows a participant to affect the environment if needed, thus, reflecting an ability to contribute.

Both our interview results and the results of the narration experiment showed that in the meeting context, functionalities of Following Dynamic Processes often revolves around being able to follow the work on a whiteboard or screen (e.g. *"he is talking to the guy who is drawing, instructing him on how to draw stuff or something like that"* [P2V6]), or to follow a specific conversation when a discussion splits: *"If somebody is talking, trying to convey something and then somebody mixed into that whispering. It's not distracting to the group, it's just distracting to you [when remote]".* [P9].

4.7 Discussion

In this project, we address the need for systematic understanding of how attention processes might be supported through video meetings features. The consolidated results of two studies on prioritization and sense-making of attention information allowed us to construct a three-level classification of functional types of observable attention in meeting communication. The classification shows that attention processes that might be accessible for the

feature support fall into one of three categories – Attention as Direction (only process), Attention as Action (process and purpose), and Attention as State (only purpose) – some of which have corresponding functions and functionalities.

The first category of visual attention that a video meetings feature might support is Attention as Direction, which simply denotes the observable information on a direction of one's gaze and pose. Attention as Direction presents a fairly straightforward entry point for developing new video meeting features, for example, due to the possibility for direct automatic recognition [345, 346, 512]. However, while this information might be used to infer that one is 'paying attention', it is considered to be not particularly informative on its own. For example, a specific instance of Attention as Direction might either reflect an 'accidental' target of averted gaze (e.g., Alice is looking at the desk while listening to Bob) or a partially perceived attention processes when there is not enough context to infer the purpose of an attention act. Thus, although this type is the simplest for technological support, Attention as Direction can not be fully considered as a functional type of attention.

Attention as Action and Attention as State, however, differ from Attention as Direction by augmenting this pure visual information with a contextually interpreted function of corresponding attention processes. The implied purpose of attention processes can be compared to the notion of a goal in goal-directed actions in the activity theory [514]. Within a three-level structural model of activity, Leontiev [270] differentiates between object-oriented activity, goal-directed actions, and basic operations. While an object is a motivation for the complex activity, goals of actions are more temporary and individually focused. Furthermore, Attention as Direction can be loosely compared to the Leontiev's level of operations that do not have their own goals [223]. For in-depth analysis on activity theory see [223, 514]. Similarly, describing pointing as situated practice, Goodwin shows that an act of pointing becomes interactive through an association with a purpose recognized in the context of the communicative situation [180]. Goodwin uses the term 'activity framework' to refer to a candidate target connected to the system of recognizable activities within which that target is embedded. This approach suggests that as well as the object that one is pointing at (as in Attention as Direction), the possible operationalized implication of this pointing act (as in Attention as Action) should be considered part of the meaningful communicative practice.

The differences between Attention as Action and Attention as State also find reflection in previous research. For example, reviewing shared attention theory, Shteynberg [430] highlights the distinction between the psychological state of shared attention and the activity of attending together with another social agent, noting that while these phenomena can affect each other, they might also exist independently. The interdependent implications of awareness as a state and of the 'shape' of actions and particular activities related

to participation in the process are also discussed by Heath et al. [200].

Furthermore, our results demonstrated that the accessibility of means for the efficient behavioural engagement (performance of Attention as Action) contributes to the participant's overall sense of engagement. We found that these effects seem to be precipitated by the limitations that remote participation puts on the three functions of visual attention.

While the Gathering Information function might be the most intuitive, the importance of the Communicative Signaling function of attention is also reflected in previous research. In Biocca's words, "social presence is not just sense of the other, but it is very much a sense of the others of 'me"' [47]. Similarly, Goodwin [179] has previously acknowledged that for an action to become social it should include not only the party producing the action, but the recognition and understanding of this action by others present in the social situation. Analyzing the literature on Social Attention, Salley and Colombo [405] offer a framework of conceptual approaches based on the functions of social attention process. They distinguish social attention as the social behavior enabling interaction with the social world, social attention as social motivation to attend to and engage with the social world, and social attention as basic visual attention to the social world. Within this framework, attention as social behaviour can be mapped to the Communicative Signaling function in our classification, as it outlines the communication aspect of attention, whereas the basic visual attention approach can be mapped to the Information Gathering function, since it refers to inert acquisition of information.

The central characteristic of the Following Dynamic Processes function is that this attention is paid to the dynamic changes within the immediate environment which consequently allows a participant to affect the environment if needed. Previous research has recognized the importance of responsiveness to other's activity within the mediated environment [47], as and the relevance of the dynamic awareness of the local environment as it often contains "a diverse and shifting display of different forms of information which are more or less relevant to the activities in which participants engage" [200]. This is partially supported in video meeting systems that enabling users to refer to and point at objects and artifacts within each other's remote environment [288, 345, 346].

4.7.1 Leveraging the Attention Classification in Design

Our classification revolves around procedural aspects of attention. The categories are not mutually exclusive and instead build upon each other (e.g. Attention as Action must include Attention as Direction). This structure aids in system design by providing a language that allows a developer to identify what type of attention and what specific capa-

bilities a particular feature might choose to enable. In other words, instead of developing video meeting features intended to foster attention 'in general', the classification allows developers to identify and design targeted features to deal with specific attention-related problems. For example, features aiming to support the Information Gathering function of Attention as Action should involve timely augmentation of information streams, e.g. narrating, zooming-in, or focusing system's view or bandwidth (e.g. [51]). Supporting the Communicative Signaling function of Attention as Action necessarily requires some form of notification about one's actions delivered to other meeting participants (e.g. [213]). Augmenting the Following Dynamic Processes function of Attention as Action should include the remote participant's ability to split system views, signal their intentions, and selectively contribute to simultaneous processes (e.g. [389]). Supporting Attention as Direction, on the other hand, would only require augmenting information on one's gaze and/or pose direction, e.g. providing gaze heatmaps or projecting gaze directions on the video stream (e.g. [420]). Finally, directly augmenting Attention as State is challenging since it does not constitute any particular process, however, once achieved, Attention as State can be supported, for example, by eliminating distractions in remote participant's environment (e.g. [211]). Furthermore, the presented classification informs how attention can be deployed in different ways by participants with different abilities in different meeting configurations.

4.8 Conclusion

In this chapter we have presented an empirically grounded functional classification of observable attention in video meetings. The classification includes categories of attention accessible for technological support, their functions in a meeting process, and meeting-related activities that correspond to these functions. We demonstrate that a particular type of attention – Attention as Action – provides special interest for technological feature development due to its structure and effects on participants' engagement. Designing a one-size-fits-all solution to 'fixing' attention asymmetries in video meetings is challenging because of the astonishing range of idiosyncratic conditions and personal needs. We hope that the granular model presented here could inform the design of systems that can capture a wide range of behaviors and correspondingly balance personalized attention processes with group attention needs that could be contextually specified.

The type of meetings we examined were work meetings between individuals. In the next chapter, replicating our research progression from part 1, we look at more intimate relationships between participants. We do this by focusing on parents of children and their use of information technology to monitor their children.

Chapter 5

Project 4: Parent's Information Needs and Acquiring Decisions in Use of Monitoring Technology

Parents feel an obligation to monitor their children to ensure their safety and well-being. In this chapter, we explore parents attitudes toward this technology, and find that its primary use-case is an information channel to better support communication and connection with their children when together. One emergent theme from this research is that the goal of these technologies should be to provide information in support of communication, not to enable constant monitoring.

5.1 Introduction

Many technologies, including diverse wearable devices such as smartwatches, fitness trackers and GPS tags, are designed for children but marketed as tools to help parents track and monitor their young children [118, 401]. However, while parents are the target consumers of these devices, the question of what parents want in terms of monitoring technologies for their children and their underlying motivations to acquire these devices are not fully understood [440], particularly in the case of parents of young children (e.g. under age 12) [45]. Specifically, questions including how often do parents want to check in, what types of information are parents interested in, how do parents use this information, what might parents' attitudes be toward various forms of monitoring technology, and what factors impact these attitudes merit further exploration.

With exploration of information needs, related motivations, and use-cases, a more general question regarding use of these technologies arises: does parental use of monitoring technologies address pure awareness needs or does it potentially go beyond awareness and toward a sense of connectedness? For example, previous research has found that new parents use information technology to keep logs of their child‘s data, and this information allows parents to get to know, understand, and synchronize with their newborns [175]. Furthermore, given a broader body of work specifically on monitoring of family members [319] and on family-oriented information systems in general [461, 276, 329, 518], we ask if wearable-based monitoring systems for parents can be viewed as part of the general spectrum of family-oriented information systems or if they should be considered as a stand alone type of system, for example, due to their unique user arrangement. Addressing these questions is especially valuable as it potentially contributes to building a theoretical understanding of core motivations of use of monitoring technologies.

Given a paucity of relevant data on monitoring technology for children [45], we first analyze information needs for parents of toddlers and school age children. Based on empirical results, we show that these needs are structured around four categories of information: *routine information on everyday activities, health information, schooling information*, and *social-emotional information*. We then explore parents' existing strategies for consuming everyday information about their children – i.e. the way that this information is leveraged. We show that the motivation for acquiring this information is a need for action on the part of the parent. Thus, by categorizing required actions, we identify three major uses for this information – *adjusting routine activities, performing emergency activities* and *building communication actions*. We also present an analysis that demonstrates that the motivation to acquire information is affected by parents‘ beliefs and values, specifically beliefs about privacy, trust-based relationships, and a need for work-life balance.

Together, these results allow us to show the need for a combination of awareness and connectedness aspects in information systems targeting parental routines. We argue that the design of systems for parental information needs should consider these needs in direct relation to the intended use of the information within the parent-child dyadic relationship. Our results suggest that current technologies marketed to parents do not directly address common use-cases for acquiring and using information. We explore our findings in the context of the literature on family-oriented information and communication systems, and demonstrate the similarities between requirements for information systems targeting parental routine and broader family-oriented information systems. Finally, we operationalize our results by summarizing specific requirements for parental information systems in a set of design recommendations to directly assist monitoring system designers in creating technologies that reflect parents' information needs and satisfy underlying motivations for

acquiring information.

5.2 Related Work

Our work was largely inspired by the rise of Child Surveillance Technologies (CSTs), and, in particular, the availability of Child GPS Technologies (CGTs). Bettany and Kerrane [45] summarize the debate surrounding these technologies. They note that literature in this domain is sparse (and that some of it is dated [440]), that studies of CSTs have frequently ignored the voice of parents, and that CGTs, despite being available, are frequently described negatively by popular media. They address this gap by leveraging analysis of internet-based data (passive netnography) to examine both the negative and positive perspectives on this specific technology. They note that attitudes toward these technologies exist at the intersection between neutralization theory [303] (where positive and negative aspects of technology are – sometimes falsely – described as equivalent) confounded with the social-technical affordances [133] and the simplicity, ease of use, and potential benefits of the technology in certain targeted situations. While past work by Bettany and Karrane [45], Marx and Steeves [303], and (indirectly) Dubbeld [133] has begun to explore the rich debate around use of child monitoring technologies, it also has limitations. Bettany and Karrane acknowledge that a significant limit of their work is methodological – passive netnography frequently provides access to strongly held opinions and should be balanced with competing methodologies. Marx and Steeves work [303] is primarily a critique of currently existing technologies and arguments against their adoption. Finally, Dubbeld's work [133] focuses on medical care for heart recipients and its generalization to parent-child dynamics remains suspect. In essence, past work has been scoped either methodologically, technologically (focused on device rather than on needs, wants, and how technology serves those needs and wants), or population-wise (focused on health monitoring or monitoring of elderly relatives [319]). The only work of which we are aware that has explored specifically the day-to-day information needs between young children and families with an eye toward technology is an exploratory study by Brown and Stockman [65]; however, their research focused specifically on how thematic analysis might help de-construct the problem of technologically mediated communication for family members with young children.

In our work we focus on understanding routine information needs of parents of young children and the potential for monitoring technology design in this domain. We first explored the literature on attitudes toward technology in parenting to account for the broader context of beliefs and practices that might affect parents' perception regarding the use of specific technology in parenting. A diverse body of research on parents' perspectives

on technology [228] reveals that these perspectives might be inconsistent for different stages of parenthood. Specifically, studies show that, for expectant and young parents, technology is used for learning [160], connecting, and distracting from the mundane [175]; at this stage, parents seem to perceive technology as valuable, fulfilling the role of mentorship that many new mothers lack as they frequently live geographically separated from family during childbirth and parenthood [175]. However, as children age, studies exploring technology in parent-child relationships become more conflicted, and the impression of technology and its value shifts negative, particularly due to the issues of work-life balance [90, 195, 204, 215, 347] and "distracted parenting" [19, 205, 368].

While previous research demonstrates mixed parental attitudes toward the place of technology in parenthood, and the increasingly conflicting sentiment as children age, one important positive role of technology is supporting low-cost awareness and social connection between family members. Early work on monitoring of family members was conducted by Mynatt et al. [319], and explored the design of a digital photo frame that helped adult children verify the condition of elderly parents living apart. Since this early work, the HCI community has explored various technology form factors [22, 375]. Early field trials [396] have frequently validated the utility of these remote family monitoring technologies.

Alongside benefits, privacy has become an increasingly important consideration, especially with respect to the introduction of ubiquitous technologies in the home [76]. Shankar et al. [423] explore the issues surrounding monitoring technology of elders and elders' privacy through the deployment of a number of prototypes. They explored reactions among monitored elderly family members to these prototypes and results indicate that health and security are valuable aspects of everyday technology, but that monitoring of activities of daily life (wake-sleep or checking medications, for example) and social connection technologies were met with greater resistance. Similarly, Tee et al. [461] argue that technology aiming to facilitate family interaction should balance awareness and privacy. Furthermore, working with extended families, authors found that parents of adult children tend to limit their presence in their child's life even when more communication is desired, as they believe that their child might not share this desire. Authors highlight the existing practices of careful considerations of family members' feelings when communicating or sharing. They show that people try to share only what they expect to be interesting to the other party and not to overload their family with information.

Much past research on information systems requirements focuses on long-term and long-distance separated families. Repeatedly, studies show that distance affects desired balance between awareness and connectedness. Neustaedter et al. [329] identified three interrelated categories of awareness information – location, activity, and status – and showed that all three types of information are desired more frequently for home inhabitants in comparison

to intimate social and extended social groups. Tee et al. [461] found that people were more interested in sharing photos with extended family living far away and calendar information with those living nearby. Similarly, those family members living far away were not particularly interested in constant updates. However, their interest in constant updates was significantly larger if there were children in the family. In a related vein, researchers have also explored technology mediated solutions for connecting with grandchildren or children over distances based on shared activities. This, for example, includes work by Raffle et al. [371] on remote reading, Follmer et al. [159] on shared play over video conferencing, and Yarosh et al. [519] on combining video conferencing with a shared tabletop display for group activities. Further, several studies on supporting close family members living apart showed that, while only a small portion of communications have social content [394], shared information on every-day experiences provides affective benefits, specifically a sense of connectedness [299], and awareness information is often appreciated as a way to start an emotionally deeper conversation [394].

The above monitoring systems exist within the domain of family-oriented information systems, and our research into CSTs/CGTs can be situated relative to past research on these systems. First, considering methodology, family-oriented information systems typically focus on awareness in the domestic realm, on family connectedness, or on a combination of both. One of our goals was to understand the environmental and motivational contexts of use of parental everyday information and, consequently, to understand the appropriate balance between awareness and connectedness that should be supported by CSTs/CGTs. For this reason, methodologically, we chose an exploratory ethnographic approach, examining motivations, use cases, and practices through experience sampling and semi-structured interviews [329, 123], as opposed to exploring adoption patterns and caused behavioral changes [299, 394] for a pre-designed system, either of the social/communicative [299, 519, 516] or informational/awareness [328, 319] type.

Previous research on family-oriented information systems has also considered a variety of family configurations. Some view the family as the household, others consider the family to be a larger microcosm. Neustaedter et al. [329] look into people's awareness gathering techniques in mixed households including teenagers, young-mid adults and middle-aged adults; Tee et al. [461] explore communication and sharing between extended families; finally, Lindley et al. [276] test a situated messaging device for home, with households that formed small networks of family and friends, and 'lone' households, who used the system in relative isolation. We, however, concentrate specifically on the role-based parent-child dyadic relationship within a family, a subset of the family unit.

Furthermore, alongside restricting ourselves to the parent-child dyadic relationship, we look into everyday short term separation (during the work/school day), as opposed

to long-distance separation, such as, for example, in research presented by Yarosh and colleagues on work-separated families [518] and long-distance parent-child interaction in divorced families [517].

5.3 Methods

The presented findings are drawn from a three-phase study, where phases were performed sequentially. Our target demographic was parents of young children, with the specific goal of understanding their information needs. The three phases for our study were as follows:

- First, we conducted experience sampling [263] to elicit in-context information from a set of parents (n=10).
- Second, we conducted a series of follow-on semi-structured interviews [496] with a subset of our experience sampling participants (n=5) and cross-validated these data with a set of new participants (n=7).
- Finally, we performed Internet-based data collection (passive netnography) [241] to further saturate our data, i.e. to address gaps in attitude via targeted analysis of publicly available data.

In the following three sections we sequentially describe the study procedures, participants, and results from each phase of our study.

5.4 Phase I: Experience Sampling (ES)

5.4.1 Study Design

ES prompts individuals to provide self-reports at random occasions (typically some small number of times per day) over a period of time, typically one to several weeks [263]. These prompted self-reports, designed to be completed quickly and in-the-moment, provide insights into daily experiences of participants. The advantage of experience sampling is that, alongside frequency data, it captures real-world behaviours "in the wild". Participants enter information during their everyday activities, reducing the likelihood that specific datum will be misremembered or not reflective of in the moment aspects of behaviour.

Participants

Our sample included 10 participants (see Table 5.1), 5 mothers and 5 fathers (age 32-50, $\mu = 39.2$). Two participants were a couple; others were the sole participant in their family. Each participant had at least one child between the ages of 3.5 and 10 years ($\mu = 6.1$). Participants who had several children were asked to focus on a single child for the duration of the study. The rational for the specific age-interval is three-fold. First, at this age interval, children are often separated from parents during the day and have their own activities during this time, such as daycare/school activities, social events, etc. Second, children of this age are not yet left unattended for extended periods of a day, and typically are not yet commuting or socializing independently. Third, children of this age are capable of verbalized communication: while children younger than 3.5 years consistently fail to deceive a competitor [435], 3-4 year olds start being able to hide information and to lie in simple settings [304]. Most children's conceptions of lies develop until about 10 or 11 years of age [477], and children's ability to flexibly vary their trust and trustworthiness behaviors based on self-interest was shown to be common among 5-, 7-, and 9-year olds. [385].

Table 5.1: Demographic information for ES participants: genders, age, number of children, duration of participation in weeks, participation in semi-structured interviews.

	P1	*P2*	*P3*	*P4*	*P5*	*P6*	*P7*	*P8*	*P9*	*P10*
Gender	F	F	F	M	M	F	M	M	M	F
Age (years old)	32	49	48	50	33	32	37	35	42	34
Number of Children	1	1	1	1	2	1	3	1	2	1
Duration of participation	3w	2w	3.5w	3w	4.5w	3w	3.5w	6w	3.5w	2w
SSI Participation	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No

Study Process

Participants were prompted 3 times a day through a specialized mobile phone application [119]. With every prompt participants were asked the following:

- Where is your child right now? [*Multiple choice*: With you; with other parent; with trusted adult, at school/daycare, apart overnight or longer];
- Describe your child's mood the last time you saw them. [*9-level Likert scale*: very upset to very excited/happy];
- How worried are you about your child right now? [*9-level Likert scale*: not worried at all to very worried];

- If we could provide you with information on your child right now, what would you like to know? [*Open ended question*].

Participation lasted 2 to 6 weeks (ave. 3.4 weeks). The mobile phone application uploaded participant responses to our central server in real time, allowing on-going data monitoring and analysis. We received 447 individual responses (272 from fathers, 175 from mothers).

5.4.2 Experience Sampling Results

Types of Information

We first analyzed the answers for the open ended question for half of our participant sample to identify a set of pattern codes, and tested these categories against the remaining participants. Two researchers conducted the initial coding, collaboratively refined the pattern codes, and the codes were then applied to the remaining data, as we continued to seek contradictory data points to verify saturation in our data. To summarize data topics, we used a descriptive coding approach based on similarity and frequency of answers. Some answers were assigned multiple codes. As a result, four general themes, or categories of information emerged in our data (see Table 5.2 for details):

- Routine information (53% answers) including general information on a child's state, such as information related to sleeping, eating, location, current activities, and school-related activities and academic progress.
- Health information (12.3% answers) including general health such as blood pressure, nutrition, and general physical wellbeing, and information related to a number of special situations, e.g., if a child is or was recently sick, this information can be related to their treatment.
- Social-Emotional information (36.5% answers) including interest in general mood and psychological wellbeing of the child. It also includes information on social interactions and personal interests.
- Nothing (23.9% answers). We specified this as a unique category due to its frequency. This includes answers where a participant indicates to be not interested in any particular information at the moment of prompting.

Table 5.2: Categories and subthemes of information parents are interested in when apart from their child with examples of answers provided during the experience sampling study

Categories and Subthemes of Information

Routine Information

Subtheme	*Examples*	*Subtheme*	*Examples*
Sleep (13)	"Did she take a good nap?" "How soon she will fall asleep" "Is he sleepy now?"	Food (67)	"what is left in his lunchbag" "Did she have enough time to get breakfast at school" "Are you hungry?"
Physical activity (81)	"How tired she is" "activity level, running, etc at recess" "If she had to use the washroom" "Is he warm enough at recess"	Activities / organization (18)	"What are they doing right now" "Did you brush your teeth yet?" "what activity is he engaged in" "is she making a mess upstairs?"
Location / safety (23)	"location. did she get to mom's classroom with mom" "Where is he?" "have they left school"	School related information (35)	"How he's doing in math class" "behavior during morning at school, listening, working, etc" "If she is doing homework now" "what is he studying at school"

Health Information

Subtheme	*Examples*	*Subtheme*	*Examples*
General well being (25)	"How are you feeling today?" "Her blood pressure" "how healthy he is" "If she has enough vitamins"	Special situations (30)	"Do you need more medicine?" "If she still has ear infection" "Why she has a headache" "If she has a fever" "still worried about her stomach"

Social-Emotional Information

Subtheme	*Examples*		*More Examples*
Mood / getting to know a child (163)	"How to help her to deal with bullies at school" "Are you worried about your surgery today?" "We need to learn more about social interactions" "How he's feeling about school today (ie sad, calm, excited)"		"Why she was sad" "Do you like playing with Adam?" "His mood" "If she likes her new dress" "What do you want to do this afternoon" "If she miss me" "How is his communication with other children going"

Table 5.3: Average levels of worry and total number of answers for each location and level of child's mood

Location	*Average Worries*	*Answers*	*Mood Level*	*Average Worries*	*Answers*
With you	1.46	172	Upset	3.6	7
With other parent	1.5	64	Some- what upset	2.8	29
With other adult	3.57	7	Calm	2	71
At school / daycare	2.16	197	Some- what happy	1.8	205
Apart over night	3.86	7	Happy / excited	1.5	135

Understanding Levels of Concern

In our analysis we found correlations between "Nothing" as information desired and level of worry, between location and level of worry, and between the child's mood and level worry.

Interestingly, the frequency of answers in the "Nothing" category seems to correspond to the results for work-separated (longer-term separation) families reported by Yarosh and Abowd [518]. Their research shows that at least some parents and the majority of children are willing to minimize long-distance interaction. However, in our results, the answer "Nothing" appeared only when a child was with the participant, with the other parent, or at daycare/school; never when a child was with "other adult" or "apart over night". Overall, 94% of "Nothing" answers correspond to the reported lowest level of worry about one's child. See Figure 5.1 and Table 5.3. Together, these results show that, first, while parents rely on a "collocated adult", as shown in Yarosh and Abowd [518], the role of the adult (parent/teacher vs. other) is a potential predictor for the level of worry; second, with higher level of worry, significantly higher desire for information is reported. At the same time fathers more often categorize their child's mood as very happy (41.2% fathers vs 13.1% mothers).

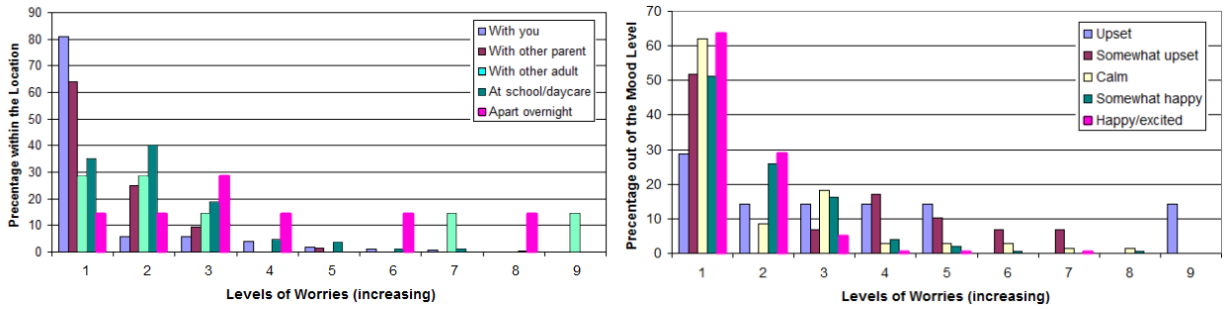

Figure 5.1: Left: Level of worries for each type of child's location. Right: Level of worries for each level of child's mood.

5.5 Phase II: Semi-Structured Interviews (SSI)

5.5.1 Study Design

The primary output of ES data is the understanding of the specific categories of needed information, specifically routine information, health-related information, social-emotional information, and no particular information as identified in Table 5.2. However, what was not yet understood is in what form parents want this information, the importance of immediacy, and what their motivations are for use of this information.

While ES method provides ecologically-contextualized data, these data are expected to lack the contextual depth of narratives, particularly because participants are prompted to respond during their everyday lives, and so responses are often short to limit the cost of interruption. Thus, ES method tends to provide replicable results that lend themselves to comparison and contrast, as well as allows to trace flows of action, but lack contextual richness of the provided narrative.

To expand on the depth of the narrative context and to capture additional nuance associated with the data – particularly how much, how often, in what form, and how important data capture is – we conducted a series of semi-structured interviews [496] (SSI) with parents of young children. In total, we interviewed twelve parents (age 32-50, ave.age 38.9) of seventeen children (age 3.5-10, $\mu = 6.1$). The participant sample had biased representation as all SSI participants had a tertiary degree (versus 63% of the Ontario population and 57% of the Canadian population [77]), 10 worked full time and 2 worked from home (see Table 5.4).

Initially, to triangulate data from ES and SSI, as well as to ensure the continuity

of analysis between different phases of the study, we conducted SSI with a set of ES participants (n=5). Aligned with recommended practices in qualitative data analysis [107], we performed incremental data analysis across participants using open and axial coding seeking saturation in our data set. Saturation occurred rather quickly; data from both ES and from SSI were highly consistent. However, analysis of data also indicated that participants had been primed by their participation in ES; in particular, participants being prompted for what they wanted to know regarding their child caused them to examine and verbalize their needs for information. To address this, and guided by the principal of selective sampling [107], we added a set of seven participants – with at least one child between the ages of 3.5 and 10 years old – to the SSI data set to ensure that data saturation was not a result of ES priming.

Table 5.4: Demographic information for SSI participants: genders, age, number of children, employment (full-time - FT, work from home - WFH), and participation in experience sampling.

	P1	*P2*	*P3*	*P4*	*P5*	*P6*	*P7*	*P8*	*P9*	*P10*	*P11*	*P12*
Gender	F	F	F	M	M	F	M	M	F	M	F	M
Age (years o.)	32	49	48	50	33	33	46	34	38	35	34	35
Children	1	1	1	1	2	2	1	1	2	2	2	1
Employment	FT	FT	FT	FT	FT	WFH	FT	FT	FT	FT	WFH	FT
ES Part.	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No

Study Process

Interviews began by asking parents what they wanted to know about their child when apart during the day. One advantage of using a smartphone application for data collection is that, for those participants who had participated in ES, we could pre-analyze and leverage the data collected from that phase of the study to seed the interview. We then drilled down on topics leveraging the freedom of the SSI format to pursue deeper rationale. We asked participants to discuss scenarios for when they wanted or did not want specific pieces of information (i.e. what motivated that information, when would they want/not want to receive it, how frequently), what were their current practices regrading gathering this information, and what they wished to do with the information. Finally, we also note that, while both ES and SSI participants' children were between 3.5 and 10 years old, during the interviews (SSI) participants discussed their children at different ages, referring both to situations in the past and in the present (if they had older or younger children).

We transcribed all interviews, performed open coding, identified information clusters, and developed overall themes from these and open coded ES data. Coding was performed primarily by the interviewer, with on-going input from a second researcher.

5.5.2 Semi-Structured Interviews Results

We started each interview by asking parents what they want to know about their child when they are apart during the day. We found that, for parents in our study, it was somewhat difficult to formulate specific questions and the type of information they wished to receive. Participants would often indicate that it was difficult or would start with general answers about happiness and wellbeing.

> *I'm not sure that I was ever clearly formulating it for myself. But in general, subconsciously I guess, I want to know that he is doing overall fine and not upset. [P5m]*

> *If they are safe... If they are happy... That kind of thing. [P9f]*

Some participants from the ES study stated that they were troubled by the question eliciting information on what they would like to know right now. However, the experience of repeatedly answering this question made some of the participants more aware of their interests and they became more able to articulate what they wanted to know.

> *These questions helped me to systematize my general emotional background, I guess. My feelings and my [pause] what I actually was interested in. I found that I was more interested in questions about her feelings and perceptions. Like why does she like this book, for example. In everyday life you usually don't really ask yourself these questions . . . You do look at facts, but you don't [pause] I mean it's almost like you don't formulate it as a question . . . [P2f]*

Resistance to Monitoring Technology

In ES data, 76% of answers provided by parents did indicate specific information of interest. However, SSI data further showed generally negative attitudes towards acquiring this information from monitoring technology. Specifically, for all types of information except nothing identified in ES, we observed a resistance to the idea of ongoing monitoring unless in special situations. We identified two primary rationales for this resistance: need for work-life-parenthood balance, and a child's right to age-appropriate levels of privacy.

First, parents expressed a necessity to limit constant information due to stress and distraction, and their need to balance work and home.

I'm a full-time working person and can't all the time think only about my child...And I think this information is not really needed. [P2f]

Similarly to Yarosh and Abowd [518] who showed that some parents rely on the collocated adult to take care of a child while they are away, and confirming results from the ES study, we observed that parents strongly believe that if there is a significant problem, then their child's caregiver will notify them.

They'll call us if he is sick, or hurt himself or something happened. . . . It's not like I don't [pause] I was almost about to say I don't really care about what he is doing during [pause] because I'm kind of busy with [work], it's not like I'm able to think all the time [about] what he is doing. [P7m]

Some parents have experienced, themselves or at second-hand, the pitfalls of monitoring and its link to stress and constant intervention. For example, several comments during SSI focused on cameras in daycare and how these cameras can encourage parents to try to constantly solve problems.

I have one friend whose child is in a daycare [with cameras]. When they joined, she was watching it all the time . . . and was constantly stressed, like "why is he sitting at that place or why he is sitting alone". So she would start calling right away: "why is it this way, why is it that way? Do this, do that". I think it is very unhealthy situation! The more details you have, the less healthy you, parent, are getting, you know. More reasons to go insane. [P1f]

We found that parents generally accept possibilities of negative experiences in their children's lives, and they cannot always be there to solve problems. The challenge for a parent then becomes knowing there is a problem but knowing they cannot do anything about it.

I think, If I just knew that at this exact moment she is crying I probably don't wanna know that. Because then I will be like "why is she upset? What's going on?" . . . I'm thinking right now she is up from her nap, she is probably playing outside, she is probably having fun. That's what I want to choose to believe. [P7m]

Bartholomew et al. [36] showed that, for new mothers, more frequent visits to Facebook accounts and more frequent content management were each associated with higher levels of parenting stress. However, general resistance to ongoing information about one's child as a mechanism to manage this stress and detailed discussion of aspects of this resistance to information and a need to believe that children are happy and not in distress have not been shown earlier.

Second, within the overall theme of privacy, there was a strong belief that children need to feel that they can avoid constant surveillance.

If I was a child, I would be bothered by the idea of being watched. I think, she needs to have her own life, her own events, that she can tell me about later if she wants. [P1f]

I wouldn't want to monitor their emotional state. It would be almost too much to know it. I think they would be feeling like I'm watching everything that they are doing. [P9f]

Further expanding on this theme, parents want to connect with their children; however, similar to extended families' households [461], parents recognize the need to balance awareness and respect for privacy. While there is a desire to know things about their children's lives, parents express the belief that they need to be respectful to limits on what they will or will not know. Parent participants recognized the ambiguities associated with their desires to know measured against their children's needs to begin to carve out space for their own experiences that might exclude their parents.

I think I would like to know if he would like to tell us . . . I don't really think that we miss out much on that part of his life except for that stuff that he chooses not to share with us which is really his privacy [pause] although, it would be nice to have a little bit more of a connection, that is kind of a fine line. [P3f]

It is better for our relationship if we, say, discuss it. Watching him would be a violation of our relationship and of trust between us. [P6f]

In essence, we see that parents express a variety of information needs while considering "too much" information to be unhealthy for themselves (level of stress, level of distraction), for their children (need to develop self-sufficiency) and for the parent-child relationship (violates privacy, violates trust in relationship). This tension in the results naturally raises a question on how information can be delivered to appropriately address informational needs but not to trigger associated resistance discussed above.

Monitoring Actions vs. Pervasive Monitoring.

Probing this issues of resistance, we note that ES data allowed us to identify three major descriptive categories of information needs (excluding nothing): routine, health, and social-emotional information. However, SSI data enriched these descriptive results by allowing us to consider a functional perspective on these information categories. We explored motivations for acquiring information and related information uses. Here we present this analysis with respect to each information category. We highlight the action-based nature of intended use of information and its importance to the format of information delivery.

While ES indicated that a large number of information needs were related to health information, SSI revealed an unexpected resistance to micro-health data and health monitoring. ES data identified two topics in health: general well being and health information in special situations, such as sickness or surgery. During the interviews most of the parents mentioned that when they were apart from their child they think about their child's general health. For example,"*Is he healthy... If he is well.*" [P8m], "*If everything is fine. Mostly in terms of health.*" [P6f], and "*I mean of course I'm interested in something general, like, if she is healthy, not hungry, not crying...*" [P1f]. However, when asked if they would like to receive information on their child's health on a regular basis, parents stated that their interest would only be for special situations, i.e., when their child is sick, to be updated on the progress of their child's sickness.

> *If she doesn't have any health problems like, you know, some children have diabetes, for example, or something like that, then I don't really need it, I don't really want to know all these things. Would be nice actually to know if she got enough nutritions that day or if she has oversugar... I'm really worried about it for some reason. But that's the only thing... Like once a day, maybe. Other things would be only for sick children, I think. [P1f]*

> *Her health... If she is sick I would be asking every day about her state and sickness progress... I wouldn't want to get anything on super regular basis. I need information about her health everyday only when she is sick. [P12m]*

We have seen little to no interest in getting general physiological data on a regular basis. Parents tend to see this information as surplus and unnecessary.

> *I just don't think we need that data at this age. I would like his doctor to take his blood pressure when he goes for a check up. I don't want him to wear a FitBit .... Microhealth data on a regular basis seems unnecessary to me. [P8m]*

We probed this idea of exceptional (e.g. when sick) versus everyday situations for other types of information. With probing, we found that, as with health information, interest in monitoring location was expressed only for special, unusual situations, while typical daily monitoring was seen as unnecessary or even harmful. Specifically, parents did want to know location:

> *I want to know he is where I expect him to be, that he is safe. ... Knowing where he is at and whom he is with. [P3f]*

However, the attitude to immediate location information ranged from indifference regarding the data to distaste for devices that constantly collected the data:

I don't think I would do it [check GPS data] often. Maybe if there is some field trip or something like that. But I don't think I would do it casually. Maybe once a week, but only if I'm bored, not that I really need it. [P11f]

Yeah, my wife actually was thinking about getting one of those GPS trackers, but I don't want to. I don't see why, we know where they are. ... There is no need for that. [P10m]

From the perspective of other routine information, while parents wanted this information, their primary interest was in summary data, typically at the end of their working day when they reunite with their child:

If she slept, for how long...It's important because if she didn't, then we need to go to bed earlier that evening. Also what she ate, if she is hungry, to understand what her dinner should be. I ask it at the end of the day usually. [P1f]

Some parents in our study, similar to P1f, used this information to adjust their schedules accordingly, while others would use it as help to rationalize their child's current state and adjust their own behavior.

The emerging view we identified was a distinction between monitoring specific actions and pervasive monitoring. This dichotomy between monitoring actions versus pervasive monitoring also arises in the interest towards social-emotional information. While parents showed a high level of interest in this information, quotes illustrate that the focus was on problem situations, not on-going monitoring:

[My child] is having troubles making new friends in the new city and that was on my mind a lot, if she is happy or not at school. Is she is finding friends to play with, if she feels safe and happy at school ... Usually I feel like I'm pretty sure that they are okay physically. It's more the emotional state that I don't know about. [P9f]

On a high level every parent wants their child to be happy. ... If he comes home and he is sad, someone is picking on him at school or something...yeah...that would be something I would wanna know. I don't know if I would monitor it, but I would want to know. [P7m]

Extending this last quote, we found that the idea of inter-personal problems was a significant concern of parents, focusing on disagreements and other conflicts. For example, parents were concerned about their children having "*a disagreement with someone*" [P6m] or "*bites, for example, from other children*" [P1f]. However, parents do not want this information in real time unless it requires specific action from them.

Okay, I'm told there is a conflict. I definitely won't run there to resolve it. I will not even call teachers at the moment. I'd better know about it postfactum. ... [P5m]

How do you usually know this? You don't. You learn it after the fact. And I probably wouldn't want to know it at the moment. I don't think school needs me monitoring [P4m]

Finally, attention to motivations for acquiring information and related uses for information categories identified through ES data reveal that parents' interest in school program and extra-curricular activities should, in fact, be identified as a separate information need category. While, previously, we categorized these questions as sub-themes of the routine information category, consideration of a functional perspective on this information showed its distinct aspects of acquiring and using practices. Specifically, the leading motivation for this type of information was to have a factual basis to be able to build more meaningful conversations with one's child and better connect with them. For example:

What did they have during the day. Because I've noticed it works better [pause] Asking some general question like "what was at school" is worse than if you already know what was there and ask something specific. Like today you were doing a project about animals, what animal did you pick? Because otherwise he, as children always do, says "we did nothing", but if I ask specific questions he starts talking about it. [P6f]

We might use it as a way to talk to him, like "oh, we saw a picture of your chemistry experiment, can you tell me more about that?" ... Usually he'll say "I don't remember"... And at camp they'll give you this nice piece of paper and they'll say "today we did this, ask him about that". I really like that. ... It's something to talk about. [P7m]

I guess more about what she is doing. Her interests, her music classes, dancing. ... What did she learn today? It helps me to build further communication with her. For example, if I know that she was doing this and that, was interested in something, was hanging out with someone, then it can be a topic for me to build a conversation. [P12m]

To summarize the desired information acquiring patterns: the SSI study showed that parents resist on-going monitoring; for typical situations, parents want summary information, often at the end of the day, unless a specific action is required from the parent. The interest in closer monitoring was identified predominantly for unique situations such as school trips (monitoring location) and sickness (monitoring recovery). Finally, the use-case for information – the fact that we see it used as a means for connecting with the child rather than as a goal of being digitally connected – motivates the desire for summary information in general and more detailed information only in specific circumstances – emergencies, sickness, or exceptional travel.

5.6 Phase III: Internet-Based Data Collection

One result from our analysis of ES and SSI data was a pervasive and strong reluctance among parents to leverage technology to monitor their children. While parents, despite the information needs, resist monitoring except in special circumstances, we simultaneously see an ever-expanding set of devices on the market presented to parents as safety devices for tracking their children. Arguably, if products that support tracking exist, there is a market for them. The concern we had, given the resistance our participants expressed to monitoring, was that our sample might have been biased by either education, geographic socio-cultural norms, or socio-economic status such that their need for monitoring differed from that of other parents [101, 282]. Thus, motivated by the principle of *selective sampling*, we wanted to expanded the collected set of opinions. Our aim with online data collection was to more fully saturate our analyses with balanced perspectives from both advocates and detractors of surveillance technologies and Child GPS Technologies in particular.

5.6.1 Study Design

To expand perspectives in our data set, we identified online platforms where discussions of monitoring technologies or ratings of monitoring devices were posted. This type of research has been dubbed *netnography*, a mash up of the words internet and ethnography. In defining the domain of netnography, Kozinets et al. [241] present two different types of netnography: active netnography, where participants engage in ethnographic practices through active participation in on-line communities; and passive netnography, where data is collected passively from discussion forms. Our analysis follows the second netnographic procedure, where quotes are collected passively from public on-line discussion boards without direct researcher participation in on-line communities.

We qualitatively examined parents' online opinions and attitudes towards Child GPS Technologies (CGTs), sleep and physical activity trackers for children, and the use of video surveillance to monitor children (nannycams, cameras in daycare) collected via public online reviews and discussions. All together we examined 143 quotes from unique user accounts from twenty three conversations from nine on-line platforms. To harvest internet data, we proceeded as follows:

1. We performed a series of on-line searches for reviews of these devices, as well as on-line discussion boards discussing the use of these devices.
2. To determine suitable sources for information, we manually qualitatively examined comments from the found platforms. Our goal was to identify information sources

discussing both the advantages and disadvantages of these technologies.

3. We then formed a data set from the suitable quotes from the examined information sources. Each quote was collected manually and selected based on including a stated opinion and related reasoning and/or an anecdotal illustration of the opinion regarding one or more of the above topics.
4. Once valance had been identified, related motivations were categorized using a system of codes based on content frequency and similarity, developed inductively.

One challenge that we faced was an ethical challenge of whether or not to include quotes verbatim. We provide sources for our data, but we do not provide actual quotes from on-line forums. Our rationale for this is that the quotes were not posted with the expectation to be reported in scientific papers, and directly quoting without participants' explicit permission, even if the data is available publicly, is potentially ethically ambiguous.

5.6.2 Internet-Based Data Collection Results

GPS tracking, CGTs: Pros and Cons

Debates on views on using GPS tracking were noticeably polarized. However, despite diversified sentiments specifically on using location tracking, a thematic analysis of reasoning to both use or resist such technology reveals convincing similarities with themes from SSI.

A common reason against using CGTs was built around the idea of unnecessary privacy invasion. One mother of a nine year old said that her son stopped using watches with a GPS tracker after he told her that "when he wears his watch, he doesn't feel like he is alone". In like manner, a significant number of parents view location tracking as an unjustified violation of the trust relationship with a child. Further, some parents talked about their own stress from either frequently imprecise data or simply from the ability to constantly monitor their child's locations, as always having to know exactly where their children were "gives them more anxiety than peace".

Those parents who advocated for the need of GPS tracking were predominantly concerned with the possibility of a child going missing. These discussions on extreme situations articulated the need "not to track, but to be *able* to track" a child. The commonly expressed desire among those parents was to be notified of unusual activity or when the device was removed. A segment of this sample of parents was specifically worried due to special family situations such as mentally unstable birth parents of adopted children

or unsupervised meetings with another parent. *Information Sources: [381], [118], [331], [422], [401], [284].*

Physical Activity and Sleeping Patterns

Unlike location tracking, we found attitudes toward fitness tracking technology to be predominantly positive. First, the awareness of a child's level of activity and quality of sleep is perceived as less invasive and less controversial than awareness of location. Furthermore, we have seen data on existing practices of inter-family competitions, comparison of sleep patterns, or the following of shared interests in sport-related activities [329]. Potentially due to this generally accepting attitude, it was significantly more challenging to identify commentary that included rationalizations for buying these devices (since there is no polarization). While parents still express concerns about privacy for their children for location tracking, those who used these trackers shared openly their children's activity data. The important observation here is that the resistance to electronic monitoring of children is not homogeneous; some monitoring data is perceived to be less sensitive and less invasive. *Information Sources: [331], [380], [284].*

Surveillance Technology

Other surveillance technologies exist for observing children, alongside CGTs and other tracking and monitoring tools. One of the major themes in discussions around the use of surveillance tools, such as nanny cameras and cameras in daycare, evolved around trust-based relationships. Corresponding to the results from SSI, the dominant reason for not using monitors was the necessity to teach a child about the concept of privacy. However, in addition to these concerns, online data revealed the topic of trust-based relationships with a caregiver. While some parents saw monitoring their child's care provider as disrespectful ("I don't need to babysit my babysitter"), others justify the use of surveillance tools via lack of trust in a caregiver.

Another recurring theme against surveillance was the need to "drop constant worries about every aspect of a child's safety"; parents holding this view argued that there is no reason to use caregiver services if they still need to constantly check up on their child.

Finally, one of the major arguments favoring the use of video cameras was the ability to feel more connected to the child. Some parents self-reported using cameras to check in on a child when missing them, the "need to remember why we are working so hard", the need to sense the presence of their children.

Anecdotaly, surveillance technologies such as cameras could be used for connectedness bi-directionaly. For example, a 3.5 year old would make her mother go to her room and act silly while she watched; and a 4 year old "feels a lot more comfortable about sleeping on her own" knowing that her parents can see her. This was the only evidence of monitoring technology used directly to support connectedness. *Information Sources: [283],[284],[381].*

Daily Activities for Building a Connection

Finally, in our earlier data we noted that one use of these technologies is to gather information for building conversations with young children. Therefore, we further explored discussions of strategies parents shared on collecting data for connection.

Parents shared struggles to build a meaningful conversation with young children and commonly articulate a need for more specific questions to elicit information about their child's day. Trying to find out what their children did during the day resulted in answers such as "things and stuff and stuff and things", or the "same thing we did yesterday". However, we see little evidence of parents using any particular technology to learn about activities in order to foster discussion. Similarly to SSI results, we saw that parents obtain this information by talking to caregivers, but often find it insufficient.

Overall, the need for information to build discussion has been observed in other research. Discussing guidelines for technologies to support parental engagement in the education of older children (e.g. teenagers), Wong-Villacres et al. [510] showed that successfully engaged parents try to access nonacademic information relevant to their child's academic success and named the lack of relevant non-academic information among four critical issues that the design of current technologies must address. *Information Sources: [381], [431].*

5.7 Discussion

Data triangulation from our three-phase study exploring parents' information needs with respect to their young children provides an enhanced picture of the rationale for information collection and use cases. While ES information yielded insight into what parents want to know, SSI provided context (why and how) for the ES data. To challenge the potential socio-cultural and socio-economic negative bias of the ES and SSI participant sample toward CSTs/CGTs, we also performed passive online data collection, revealing similar themes, particularly on motivations and beliefs around the use of monitoring technology

for young children. In this section, we discuss three synthesized data that cross-cut our three-phase study: information categories, information use, and connectedness.

First, based on ES results, we categorized parental information interests into three themes: routine (including schooling), health, and social-emotional information (see Table 5.2). Further analysis from SSI data showed that parents perceive and treat school information differently (see sec. 5.2.2), yielding four overall information categories: routine information, health information, school information, and social-emotional information. Each category differs in strategies for seeking the information, desired regularity of the information, and ways information was used. Neustaedter et al. [329], when studying interpersonal awareness in households, also explored information dynamics and showed that people want to know information that typically falls into three interrelated categories – location ('where are you going'), activity ('what are you doing?'), and status ('how are you doing?'). They also showed that all three types of information are desired more frequently for "home inhabitants" – the people with whom one lives. While the structure of information needs of parents of young children that we presented covers similar types of information found by Neustaedter et al. [329], it is clustered differently (see Table 5.2). For example, location awareness is predominantly viewed as a predictor of safety, not as information separated from other routine information such as food, physical activities, happenings, etc. Similarly, status awareness is clearly separated into health status information and social-emotional information categories, where these types of information are both acquired and treated differently by our parent participants. Variations in the information clustering in our results and in results reported by Neustaedter et al. [329] could be explained through the differences in participants' goals of acquiring this information and the intended use.

Extending the idea of intended information use, our results provide evidence that the information is predominantly desired when an action is required from a parent, it is used to guide this action, and it tends to be perceived as supererogatory or distractive otherwise. In his seminal paper on understanding and using context, Dey [121] gives a brief analysis of existing definitions of context. He provides examples of a definition by Schilit et al., where context includes information on where you are, who you are with, and what resources are nearby [409], and a definition by Pascoe, who understands context as the subset of physical and conceptual states of interest to a particular entity [351]. To overcome the dependency on situational aspects, Dey suggests to view context as "any information that can be used to characterize the situation of an entity", and defines a context-aware system as a system that "uses context to provide relevant information or service to the user, where relevancy depends on the user's task" [121]. Then, the information needs of parents as elicited can be categorized as context information based on the combination of the types of information (location, activity, status and health) [329] and the key role that the intended use of

information plays in the patterns of acquiring this information [121]. Consequently, the appropriate delivery of this information to parents through information systems requires we understand the user's task (the information use-cases). Our data shows that use-cases for the parental information of interest fall into one of three categories: *adjusting routine activities* (e.g., feeding a child, adjusting sleeping schedule), *performing emergency activities* (e.g., emergency pick-up from school, finding a lost child), and *building communication and deeper personal relationship with a child*. Considering parents information needs in light of the intended use seems to be a promising approach to address the tension between resistance to pervasive monitoring versus the need to monitor to identify required actions.

Finally, one of the identified categories of use-cases – that of building connection – coupled with parental privacy concerns and concerns of supporting trust-based relationships with a child, suggest that information systems to track and/or monitor children should be designed to address a combination of connectedness and awareness needs as opposed to purely supporting awareness. Correspondingly, our findings on the structure of parents' information needs, underlying motivations, and existing practices show that there are similarities in the requirements for parental information systems and a broader spectrum of family-oriented information systems. For example, we see that this awareness information is largely used to support coordination of routine activities [328] and light-weight communication [299] built around routine artifacts [164, 461]. We see that parents desire the awareness information as a means (sometimes the only accessible means) to better understand and to build a conversation with a child. Similarly, Romero et al. [394] noted the importance of utility-oriented topics as a vehicle for families to engage in deeper emotional communication. Lastly, we find that privacy concerns, previously shown for older family members [423, 461], still hold true in parent-child relationships.

5.7.1 Design Recommendations

Considering the above synthesis themes of information categories, information use, and supporting connectedness, in this section we examine how these themes can be applied to the design of information technology. Specifically, we leverage these three themes in a set of design recommendations that can serve to operationalize our findings.

Information Categories and Information Format

One distinct characteristics of information systems for parent-child relationship is that, unlike older family members, young children are not expected to provide trustworthy and

complete information when needed. Hence, it is essential for design to incorporate not only means of providing information, but the appropriately communicated information itself:

DR1: *Design to deliver information in logical clusters according to parents' information needs*

We found that some parents might have difficulties articulating specific types of information they are concerned with when apart from their child. Additionally, we have seen that parents tend to think about this information as amorphous concepts ("happiness", "what is happening in their life", etc.). We compound parents' information needs into four primary categories: routine information (i.e., location, eating, sleeping), health information, daily activities (to build conversations), and social-emotional information (to understand and connect with a child).

DR2: *Focus on delivering summarized information*

Our results show that on a typical day parents don't want the information available to them constantly as a stream of information on what is happening right now. Instead parents prefer summary information, a recounting of what has happened during the day apart. Parents were interested in getting this information, for example, before they pick up their child from daycare. They then would use this information to adjust their behavior when with a child later that day, or as a basis for shared activities.

DR3: *Deliver information when direct action is required*

The information is generally desired when an action is required from a parent, and tend to be perceived as supererogatory or distractive otherwise. As part of identified significant resistance to receiving real-time monitoring data, parents expressed a strong need to believe that if something happens that requires a specific action from them, they will be notified. The actions-based information use-cases include adjusting routine activities, performing emergency activities, and building communication with a child.

DR4: *Distinguish routine and exceptional circumstances*

With the acknowledgement of general resistance to real-time monitoring, monitoring is considered to be acceptable or even necessary in exceptional circumstances such as school trips, health recovery, chronic illnesses, or highly unique situations such as unsupervised parental visits with limited trust between parents.

Information Use and Context

Parents' resistance to pervasive monitoring of a child arose both because of concerns about children's privacy, and concerns about their own life balance and mental health. Together,

this analysis leads to the following two design implications:

DR5: *Design with consideration of the value of work-life balance*

Effectively designed technology needs to be aware of parents' need to balance other obligations – such as work – with their role as parents. Specifically, there is a dichotomy between knowledge (an awareness of a fact) and monitoring (the stream of data that supports that awareness), and technology should limit information streams to that necessary to support the knowledge of a child's welfare rather than its constant monitoring. Stated another way, designers of monitoring technology need to support parents' non-parental obligations, e.g. by prioritizing information to inspire confidence in the safety and happiness of children while limiting distractions at work such that parents can be productive.

DR6: *Design with consideration of children's privacy*

Collecting information about a child through technology was often associated with violating a child's privacy and creating a sense of "being watched". This held true for all types of information, but specifically for social-emotional information that parents mainly viewed as child's personal space. While parents have emphasized their deep and predominant interest in this information, a child was perceived as the main owner of such information. Additionally, there are developing social attitudes associated with technology-assisted monitoring. For example, heated debates in the media create a notion of social shaming for monitoring, calling parents who use monitoring devices "creepy" and "paranoid" [45].

Designing to Support Connection

The emotional bonding that family members have toward one another, also known as family cohesion [111], heavily relies on a positive involvement of parents with their children, e.g. shared activities [343], that have been linked to fewer behavior problems among children [209]. We advocate that information technology can be design to foster family connections:

DR7: *Design to support building the parent-child communication*

Parents articulate the need in an information store that notes happenings of importance in their child's day; noting that their children are young and cannot always articulate happenings; and noting that young children often believe that parents are omniscient. While, for example, wearable devices can be used to monitor information in real time, they can also be used to provide a store of knowledge on events of interest for later discussion. What did the child do in class? Was there anything special that happened?

DR8: *Design to include other caregivers as a source of information*

Parents seek information about their child's day from various sources – teachers, babysitters, children themselves. However, they articulate a gap in this knowledge, wishing that they would be provide with more detailed daily information. While technology can be used as a data store to build connection (DR7), we argue that other caregivers could all serve as secondary users of these technologies, allowing simplified sharing of information.

5.7.2 Methodological Limitations

We performed a three-phase study using experience sampling, then semi-structured interviews, and finally internet-based data collection (passive netnography). The advantages of this three-phase study are two-fold: it both permits us to triangulate data cross sources and to address methodological limitations of one study with another. As an example of how limitations are addressed, high replicability but lack of contextual richness of the narrative in ES data is balanced with the data from semi-structured interviews, which tend to produce results that are well-suited to developing thematic glosses and evaluations but might lack granularity, and thematic prioritization in self-reported information tends to relate to what is remembered. Another example is that the perceptions of CSTs/CGTs – which risk missing diversity of experience in a specific group of participants – can be countered through a broader internet-based data collection.

However, we acknowledge some remaining limitations of the methods. For example, while there is a legitimate balance between a parent's right to know and a child's right for privacy, the reflection resulting from both experience sampling and interviews can potentially over-emphasize the tension around these issues due to the explicit prompting of what one wants to know and when and how frequently one wants to know. Furthermore, while internet-based data collection can counter-balance one-sided views, the views found on-line tend to contain strongly held opinions. Thus, while the issues and tensions explored in this project are legitimate, the intensity of the effect of discovered issues and opinions might potentially be augmented through both direct inquiry with participants and because any counter to this from on-line sources allows little calibration of the intensity of opinion. Consequently, further research is required for the additional validation of this intensity.

Alongside this, one concern we had is that the one-sided views that were identified in experience sampling and semi-structured interviews might result from demographic homogeneity: our ES/SSI participants were drawn from a relatively narrow geographic region and had similar education levels. While internet-based data collection was motivated by a desire to address this issue, the method chosen for the on-line data collection does not allow access to the demographic information of posters. As a result, care must be taken in

the degree of generalization of our results. Follow-up work with different cultural groups presents interesting opportunities for future work.

5.8 Conclusion

In this chapter, we present results collected through three sequential studies: an experience sampling study, semi-structured interviews, and netnography. We first show the structure of parents' information needs with regard to their children, organized into four information categories: routine information, health information, school information, and social-emotional information. We analyzed parents' motivations for seeking this information and demonstrated how these motivations are affected by parents' values about privacy, trust-based relationships, and a need for work-life balance. We showed that the information is predominantly desired when an action is required from a parent, and identifies the common use-cases for this information. Finally, we highlight the unique aspects of requirements for parental everyday information systems via a set of design recommendations.

In the two projects described in this part of the thesis, we looked at technologies that could serve to assist people in their communication with one another. However, frequently the goal of communication is not to communicate with another person; instead, it is to communicate with technology. In the next part of this thesis, we look at this communication with technology with the goal of how users perceive conversational agents they communicate with and what the role of these conversational agents should be.

Part III

COMMUNICATION WITH TECHNOLOGY

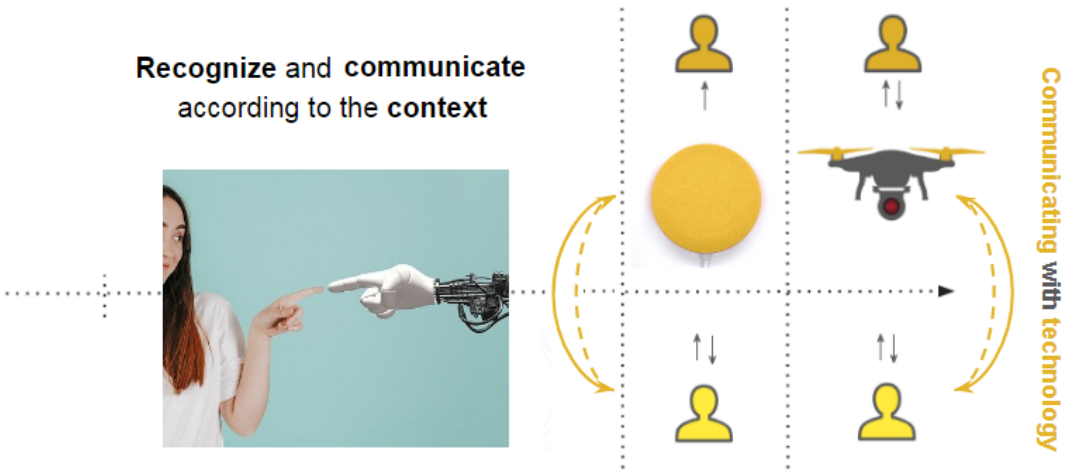

Figure 5.2: Pattern of operations adaptation in communication with the system

Lastly, we explore communication directly *with* the system. Previous research demonstrates that such communication requires a unique set of considerations of user's predispositions and biases. For instance, "automation bias" describes people's tendency to see greater authority in automation compared to human experts [350, 116], and which leads to greater perceived reliability and user's reduced information verification behavior [28]. Furthermore, although interpersonal and human-machine trust processes seem to be fundamentally different [294, 137], the media-equation hypothesis proposes that among other objects people specifically tend to apply social rules and expectations to computers [326, 323]. While users generally tend to perceive computers as human-like actors [324], their tendency for anthropomorphization (attributing lifelike qualities [145]) of the systems is often further amplified by purposefully designed anthropomorphic features [229, 230, 388] and system integration in user's social space [365, 31, 478], affecting the perceived approachibility [62, 335, 46, 285] and level of user's trust [490, 116, 365, 109, 267] in the system.

Grounded in the results from **research projects 5 and 6**, we suggest that when communication occurs **with** the system, the role of this system is to recognize the user's communication context and *to "match" this context in handling the user's information.*

Given that anthropomorphism has been shown to affect users' interactions, expectations, enjoyment, and trust in the system, **research project 5** (chapter 6) explores and identifies patterns in users' anthropomorphized perceptions of conversational agents and their association with specific designed features of the systems.

Our results are drawn from a comparative analysis of three conversational agents –

Alexa, Google Assistant, and Siri. We gathered user's perceptions of agent through a qualitative multi-phase study, leveraging semi-structured interviews primed with the interaction session with three agents, and followed by a visualization exercise. During the study sessions, participants were asked to compare specifics of each agent, and then explicitly asked to imagine and describe each of the agents as a human character. The accumulated findings from this multi-phase study demonstrate that the anthropomorphized behavioural perceptions and visual conceptions of conversational agents yield structural consistency. Furthermore, these anthropomorphized perceptions are interdependant and linked with the designed system features – the character perception is self-reportedly motivated only by the design, while visualizations are motivated by some design features, but are predominantly associated with the character's perceptions. The results from this project reveal the richness of users' socially meaningful interpretations of agent's actions, including users' extreme tendency for displaying biases and social stereotypes.

Continuing our investigation of user's interactions with conversational agents, in **project 6** (chapter 7) we explore user envisionment of the conversational agents' possible tasks, related performance expectations, and associated users' concerns regarding the involved resource dynamics.

We capture and systematize user envisionment of conversational agents' performance expectations, elicited from our participants through a probe-based envisionment exercise in the form of a thought experiment built around a hypothetical scenario. This methodology allowed us to benefit from a user-centred design approach, while exploring participants' needs and perceptions beyond their direct system familiarity. We present our findings as a four component conceptual model of the conversational agent work process. The first component is a classification of interaction types, which represent high-level functional differences between clusters of agent tasks. These types were revealed through an inductive analysis of task examples. We consider them "functional", because these differences reflect the task performance processes through the combination of user's involvement in the primary activity and the nature of interacting entities. They are "high-level" differences, because they occur at a higher level of generalization than the specific task activities. The second component of the model, the typology of envisioned conversational agents tasks, is organized according to the interaction types they represent. If interaction types reflect the high-level functional differences, the typology of tasks deepens the detailed differences. The third component of the model is a schematic representation of two task execution flows associated with parent interaction types from the perspective of resource dynamics and task outcomes. Finally, the fourth component of the model is a classification of users' concerns mapped onto the task execution flows based on the nature of concerns and their correspondence to system architecture and design components.

Consolidated findings from projects 5 and 6 reveals the importance of considering user's perceptions of a system as a communication partner when they share information while interacting directly with the system. In other words, while in the first two types of interactions (Part I and Part II of this thesis), a system acts to support one's interpretation of information based on user's initial sharing intent, in this case, a system needs to directly act itself on this intent. Thus, when communication occurs directly **with the system**, the role of this system is to recognize user's information sharing context and *to "match" this context in communication with the user*. For instance, the depth and richness of social interpretations of conversational agents, demonstrated in **project 5**, highlights the necessity to access "social context" for the interaction, as implied by the user, since user forms the sharing intent accordingly. For example, perception of an agent as "cozy and homey" can be expected to provoke different information sharing patterns compared to a "distant, cold, strictly professional" perception. Additionally, **project 6**, in turn, shows the potential for a system to appropriately match its feedback to the user's contextual expectations of recourse dynamics and task performance. Such feedback, then, would represent a communicative feedback from the system as information recipient to the user as information sender, becoming part of a continuous communication loop and re-enforcing user's information interpretation expectations [227, 34].

The following two chapters present a detailed discussion of **project 5** and **project 6**.

Chapter 6

Project 5: Anthropomorphized Perceptions of Conversational Agents

As conversational agents become an ever-more-pervasive part of our world, we often begin to view them as having anthropomorphic features. In this chapter, we explore this anthropomorphization, seeking to understand how vivid or complete users perceptions of these agents may be.

6.1 Introduction

Due to a growing ecosystem of technologies that support conversational capabilities, more and more users are interacting with conversational agents on a daily basis [85], including such intelligent personal assistants as Siri, Google Assistant, and Amazon Alexa [70]. The devices supporting conversational agents, such as smartphones, tablets, and smart speakers, are often intended to be located in a particular environment (e.g. Google Home, Amazon Echo), and rely on users' voice interactions with a persona (an agent), who is deliberately designed to speak and act human-like. These characteristics integrate assistants in both technological and social spaces [478] of a user's life, e.g. by increasing perceptions of social presence [365], and encourage anthropomorphism – attributing lifelike qualities [145] – to conversational agents.

While anthropomorphism has been shown to affect users' interactions, expectations, and overall satisfaction with conversational agents [363, 170, 224], it also introduces a number of unique adoption characteristics and integration challenges. For instance, there

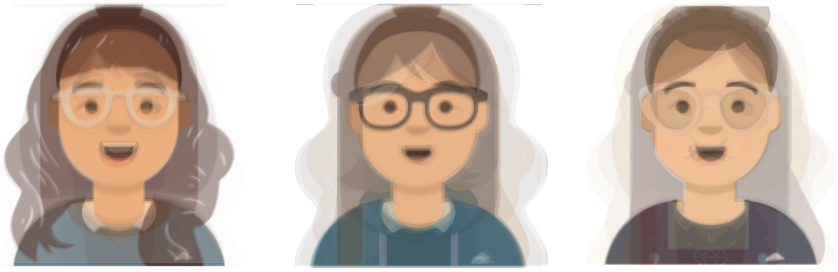

Figure 6.1: Cumulative visualizations. Left to right: Alexa, Google, Siri

is evidence that the magnitude of anthropomorphization of agents and their social behavior shapes users' trust in the system, but anthropomorphic features can also increase trust resistance [116]. Furthermore, research suggests that users' satisfaction and relationship with intelligent agents are mediated by perceived trust and enjoyment during interactions [267], which are sensitive, among other factors, to the agent's interaction style [46]. Respectively, the interaction experiences with conversational agents might potentially be more important to users' satisfaction than the actual interaction output [285]. Thus, the understanding of processes that define the structure of anthropomorphized perceptions of conversational agents is essential for appropriately designing and addressing integration challenges of this technology. However, while research actively explores the degree of anthropomorphization of conversational agents in casual use [286, 170], and factors affecting the manifestation of anthropomorphism [257, 169, 363, 146], little is yet known on the structure and mechanisms of specifically anthropomorphized perception of these agents.

To address this gap, we focus on the structure and associative relationships within the anthropomorphized perceptions of conversational agents. Specifically, we examine the anthropomorphized perceptions of three popular agents to elicit patterns of perceptions through comparative analysis. Our results are drawn from a qualitative multi-phase study, leveraging semi-structured interviews primed with the interaction session with three agents, and followed by a visualization exercise. During the study sessions, participants were asked to compare specifics of each agent, and then explicitly asked to imagine and describe each of the agents as a human character.

Research on anthropomorphic design, most developed for artificial agents and robotics [229, 230, 388], differentiates [134] between design of physical aspects of automatic agents (e.g. facial and body expressions [230, 491]) and design of their behavior traits. The behavioral aspects have been defined [198] as common and basic traits of human beings, including emotional responsiveness, interpersonal warmth, cognitive openness, individuality, and depth. By design, voice assistants are capable of displaying mainly behavioral anthropomorphic aspects, with the exception of voice, which can arguably be categorized

as a physical trait. However, we hypothesized that behavioral anthropomorphized perceptions of conversational agents would potentially affect users' corresponding anthropomorphized visualizations, especially since an agent's answers might include visual references, e.g. "I'd blush if I could" [498]. Then, if there are patterns and consistent differences in behavioral perceptions of agents, we can also expect consistent differences in related visual conceptualizations of agents. To explore this hypothesis, we started by separately examining participants' descriptions of the anthropomorphized behavioral traits and the visual appearance of the agents.

We first demonstrate that the perceptions of an agent's behavior and character are structured according to five categories, where each category represents a comparative scale: the approachability category (emotionally/socially approachable versus distant), sentiment toward a user (genuine and caring versus disingenuous and cunning), professionalism (degree of manifestation), intelligence (positive or negative), and individuality (neutral versus defined personality). Respectively, we discuss the differences in the perceptions of three studied agents within these categories. We then present the results of visualizations of the agents' physical appearance, including body build and age, facial expressions, hair, and clothing style, describe the observed tendencies in the structure of descriptions in each of these categories, and discuss the consistent specifics assigned to each agent.

These results reveal the patterns appearing within both anthropomorphized behavioral and visual perceptions of conversational agents. To further validate and operationalize these findings, we then explore the associative relationships between different descriptive elements. Inspired by the concept of dynamic anthropomorphism [269] in robotics, we consider anthropomorphism of conversational agents as two-fold. The first level is represented by intended anthropomorphic design of an agent – explicitly built features, such as voice characteristics, richness of responses, or interaction style (e.g. jokes, laughter, answers to personal questions, etc.). The second level encompasses the actual process of anthropomorphization of an agent by a user, which emerges from the interaction. This refers to the anthropomorphized perceptions developed based on users' interpretations of designed features and further attribution of corresponding human characteristics, for example, being "patient", "genuine", "stressed", "motherly", "chill", etc. (examples from our data).

To gain insight into the relationship between anthropomorphized perceptions of agents behavior and visual appearance, and their association to system features, we analysed participants' self-reported explanations and associations between these types of elements. We discuss the specifics of these relationships, and demonstrate that the character perception is self-reportedly motivated only by the design, while visualizations are motivated by some design features, but are predominantly associated with the character's perceptions.

In the remainder of this chapter, we first provide an overview of relevant research on anthropomorphism, followed by the multi-phase study design and data analysis. After presenting our findings, we discuss them in the context of previous research and potential implications for conversational agent design.

6.2 Related Work

Anthropomorphism, also referred to as personification, can be defined as people's tendency to attribute lifelike qualities, characteristics, or mental states to objects and other non-life artifacts [142, 151, 135, 490]. The media equation paradigm suggests that among other objects people specifically tend to apply social expectations to computers [326, 323], and "mindlessly" perceive computers as human-like actors, even when no explicit anthropomorphic features are involved [324].

However, a number of computational systems, including intelligent assistants, purposefully design anthropomorphic characteristics, particularly because anthropomorphism has been shown to ease user interactions. For example, Breazel et al. [62] showed that robots and other agents are more approachable and engaging when they exhibit human characteristics. Karjalainen et al. [224] studied the design space for drone companions and found that people were particularly positive towards a drone companion that featured anthropomorphic features. In a related vein, the emotion analysis of Amazon Echo reviews showed a link between personification of Echo and more positive emotions towards the agent, compared to those users who did not personify the device [169]. Earlier, Qiu and Benbasat [365] showed that using humanoid embodiment and human voice-based communication might enhance users' trust beliefs. Furthermore, Waytz et al. [490] demonstrated that individual differences in anthropomorphism predict the degree of moral care and concern afforded to an artificial agent, the amount of responsibility and trust placed on an agent, and the extent to which an agent serves as a source of social influence.

Despite the active development of the domain, the processes underlying the manifestation of anthropomorphization are still poorly understood. Some research suggests that anthropomorphized perceptions are binary and occur above a specific threshold, below which the agent is perceived as a computer [54, 116, 502]. Approaching this threshold can cause significant decreases in liking and trust around the technology displaying anthropomorphic features [292, 291]. On the other hand, there is evidence challenging this theory. In particular, Thompson et al. [467] found that ratings of the "humanness", familiarity, and eeriness of walking avatars changed continually, suggesting that users' cognitive response to anthropomorphic features of a system might develop linearly, as opposed to

occurring after a particular threshold. Finally, the model of the dynamics of anthropomorphism in robotics [269] distinguishes between the anthropomorphic features designed within a system and anthropomorphism as a process of attributing lifelike qualities as prompted by interaction with this system. Correspondingly, the model proposes to view anthropomorphization as a dynamic process following three stages: initialization, when anthropomorphization tends to increase; familiarization, associated with a decrease of anthropomorphic effects due to the user's acquired ability to predict the system's behavior; and stabilization, when anthropomorphic effects stabilize over a longer time. The measurements of a user's emotional response when interacting with conversational agents also suggest its dynamic nature during interaction stages [53].

While conversational agents include a number of explicit anthropomorphic features, findings on the actual degree of their casual anthropomorphization yield ambiguous results. For example, the analysis of user reviews of Amazon Echo showed that a significant number of reviewers were personifying Alexa as an assistant, a friend, or a family member [170]. Concurrently, another study explored the anthropomorphization of Alexa through a diary study, and found that less than half of study participants reported personification of the agent [286].

Nevertheless, research is actively revealing factors affecting the manifestation of anthropomorphization of intelligent agents. Among these factors, a humanized agent's name seems to be linked to higher personification [364], by allowing users to form more human-level relationships [169] with an agent, and to give them a social role [257]. Anthropomorphization might also occur to address the interaction contextual factors. In particular, it might help users to make sense of an unfamiliar situation when their understanding of a system is poorer than that of humans, to reduce feelings of insecurity when they seek to explain system behavior, or to establish lacking social connections [143, 142]. On the other hand, greater personification was demonstrated to co-occur with more social interactions with an agent. For instance, Purington et al. [364] demonstrated that people from multiple-member households are more likely to personify Amazon Echo than people who live alone. Research in robotics also suggests that anthropomorphization might be affected by the level of similarity between an artificial agent and a user. Specifically, Eyssel et al. [146] found that participants tend to feel closer and anthropomorphize more those robots who speak with the same gender voice as them.

Investigations of the specifics of anthropomorphized perceptions are, nevertheless, rather fragmentary. For instance, a number of studies have demonstrated that people tend to transfer social stereotypes, such as ethnicity [360] and gender biases [325, 59, 60, 147] onto anthropomorphized systems. It was also shown that the perception of an agent's personality and helpfulness has strong associations with the task performed with the system

(carrying out procedures versus providing opinions), more so than with the agent's appearance [84]. Yet, little is known about the variations of anthropomorphized perceptions. Given a paucity of relevant data on anthropomorphized perceptions, this project explores the structure and specifics of behavioral and visual conceptions of conversational agents, as well as related associative relationships with designed features.

6.3 Study Design

Our results are drawn from a qualitative multi-phase exploratory study seeking to identify potential patterns in users' anthropomorphized perceptions and conceptions of conversational agents. The study probes these perceptions by leveraging semi-structured interviews, primed with an interaction session and followed by a visualization exercise. Each study session was audio recorded with the participant's consent.

6.3.1 Participants

In total, our sample included 20 participants (age 20-27, av.age. 23), 10 female and 10 male, with diverse levels of familiarity with the targeted agents. The rationale for the balanced self-reported gender representation in our sample is motivated by previous research, which suggests that users might feel psychologically closer to agents of the same gender as them [146], and thus we wanted to balance potential impact of participants' gender on our results. Furthermore, there is evidence that perceptions and interactions with conversational agents might be influenced by gender stereotypes [59, 60, 147]. While some participants did not own a device with an agent (Alexa n=17; Google Assistant n=10; Siri n=8), others owned such devices, but had never used the agent (Alexa n=1; Google Assistant n=3; Siri n=5). Two participants (P15, P20) did not own any of the three devices, while others owned at least one. Finally, the majority of those participants who regularly use Alexa (n=2), Google Assistant (n=7), or Siri (n=7) use the default voice setting (female, American). As an exception, P3 (male) uses a British female voice setting for Siri, and P1 (male) and P19 (female) use an American male voice setting for Siri. Our rationale for recruiting participants with different levels of familiarity with each agent is two-fold: we wanted to capture the differences in the anthropomorphized perceptions rooted in experienced versus novice in-situ interactions, as well as see the effects of direct in-situ comparisons of agents on existing preconceptions. To address potential inequalities in participants' experiences with specific agents, we primed each interview and visualization exercise with a semi-guided interaction session, as described in the study process section.

6.3.2 Study Process

We started each one-on-one session by asking participants to complete a questionnaire to collect their demographic data and capture their usage patterns with each conversational agent.

We then invited participants to engage in a semi-guided interaction session with conversational agents, using an Amazon Echo Dot (3rd Gen) Smart Speaker equipped with Alexa, a Google Home Mini smart home speaker powered by Google Assistant, and an Apple iPhone XS powered by Siri. Participants were provided with three groups of three conversational prompts (access to knowledge, intellect, and personality groups), and were instructed to choose and execute one prompt per group for each conversational agent. The prompts were motivated by the results from Sciuto et al. [415], showing that users tested Alexa's capabilities by asking questions from three categories: access to knowledge – factual questions, intellect – sophistication of accessible commands, and personality – agent's responses to non-factual prompts. The methodological benefits of the interaction session in-situ comparisons of agents are such that later we were able to tease out differences as they were seen by participants. Furthermore, the mechanisms of forming a reliable perception of an agent remain unexplored; however, there is evidence that it depends on performed tasks (e.g. [84]). This motivated us to structure the interaction sessions based on previously reported types of questions used to test an agent's capabilities [415].

The interaction session was followed by a semi-structured interview. Researchers would first prompt a participant to discuss their experience with each agent, and to describe and compare aspects, specific to each agent. Participants were then asked to imagine each of the agents as a human character and to describe what their definitive characteristics would be. Consequently, and priming the visualization exercise, we asked participants to describe how each agent would look like as a human character. During the interviews, participants would often re-initiate interactions, prompting several agents with the same command to better inform their descriptions.

In the next phase – a visualization exercise – participants were invited to create an approximate visual representation of Alexa, Google Assistant, and Siri using an open source web avatar generator[1]. We acknowledge that using predefined software for visualizations might introduce a number of potential restrictions to the visualization nuances. However, we chose to use a standardized platform to overcome participants' potential resistance to a hand drawing exercise, as well as to allow better comparison of the visualization outcomes between participants, regardless of their drawing skills. We also conducted the visualization

[1]Avataaars, open source software, https://getavataaars.com/

exercise after interviewing participants about agents' appearance to ensure that limitations of the software would not affect their initial perceptions. Finally, we deliberately chose an avatar generator that had a broad and diverse representation of appearances.

We concluded with a debriefing interview to capture participants' general notes and comments on their study experience.

6.3.3 Data Preparation and Analysis

After fully transcribing audio recordings of each session (interaction sessions, interviews, visualization exercises, and debrief), we composed three data corpora for further analysis.

The first corpus included descriptions of agents' behavior and character, e.g. *"she has an authoritative vibe, but still friendly"* [P2], and originally consisted of 155 quotes. However, most of these quotes included a host of descriptive elements. For example, the quote *"Google is more genuine, and Siri is more transactional. Google is like more trusting. If it was a social setting, I would rather hang out with Google instead of Siri"* [P1] includes the following description elements – "genuine", "transactional", "trusting", "would hang out with", "would rather not hang out with". Thus, next, we extracted all the specific descriptive elements for each agent for the analysis. While some elements consisted of one-two words (as above), others were left longer (e.g. *"It's jarring when she makes a joke because it seems like she's pretending"* [P4]) to preserve semantic wholeness for thematic analysis. This resulted in 284 descriptions (Google n=102; Alexa n=97; Siri n=85).

The second corpus included descriptions of agents' appearance (186 initial quotes), e.g. *"I picture her hair as down, black or dark brown"* [P11]. We split the descriptions in these quotes (n=254) into one of the following: clothing style, e.g. *"hippy, colourful dress, really bright, green-ish colour"* [P16] (n=81); hair style, e.g. *"she always curls it, you know those wavy ones"* [P7] (n=48); hair color, e.g. *"bleach blond"* [P5] (n=44); age, e.g. *"Alexa is a younger person and Siri is her mom or something"* [P4] (n=40); build, e.g. *"Siri would be the smallest, Google the tallest, Alexa in-between"* [P10] (n=21), and facial expressions, e.g. *"a less friendly expression"* [P8] (n=20).

Finally, the third corpus of data included quotes with associative relationships between descriptive elements (e.g. *"Just seems stressed – they've got a lot of stuff and services, so their hair is a bit dishevelled"* [P7]). We then identified associative pairs of descriptive elements for each quote in this data corpus (e.g. "stressed – have a lot of stuff and services", "hair is dishevelled – stressed") so that the first element in the pair corresponds to the result, and the second corresponds to the reason ("A because of B"). We split the

resulting 167 associative pairs into either character (n=65) or visual appearance (n=102) descriptions based on the first element in the pair.

Note that quotes, combining descriptions of character, appearance, and/or descriptive association (e.g. *"taller because she is daunting"* [P8]), appeared in several data corpora.

After the preparation of data was completed, we coded [88] descriptive elements in each corpus. We first performed open coding to identify themes in the first data corpus across all three agents. We then applied the resulting coding scheme to data on each agent separately for validation and to allow comparison between agents. Each phase of coding was initially performed by authors independently, then codes were discussed and refactored in consultation with the research team. The second corpus of data was analysed in a similar manner. The third corpus of data included associative pairs of character or visual appearance descriptions, and thus, was analysed using the coding scheme from the first two data corpora.

The visualization exercise outcomes were composed into collages (Fig.6.3), and appearances of chosen visual elements (Tab.6.2) were calculated. We also composed a cumulative visualization for each agent (Fig.6.1) by overlaying avatars: avatars were split based on the hair style, and overlaid from the least to the most represented style, gradually giving those layers higher to lower opacity respectively. The results were analyzed in comparison to the second corpus of verbal descriptions data.

6.4 Results

In this section, we present our findings on the structure and specifics of behavioral and visual anthropomorphized perceptions of agents, followed by the analysis of the associative relationships between these perceptions, designed features, and perceptions of the corporate entity that produced the agent.

6.4.1 Character and Personality

We started by exploring the descriptions of agents' character and behavioral specifics[2]. We first analyzed the whole corpus of data, combining descriptions of all three agents, and identified five major themes, each representing a comparative scale for an agent's

[2]Note: The terminology presented is elicited from interview data. The associated views and opinions are those of the participants and do not necessarily reflect the opinions and vocabulary of the authors.

character. These themes include the agent's social approachability, projected sentiment toward the user, professionalism, intelligence, and specifics of the agent's individuality. We then applied the same coding scheme to the behavioral descriptions of each agent separately (Alexa n=97; Google n=102; Siri n=85), to validate these themes and to explore the differences in the assessment of the agents (Fig.6.2).

Approachability

The theme of approachability reflects an assessment of an agent as either emotionally/socially distant or approachable. The "distance" end of this scale includes such assessments, as, for example, an agent being *"less of my friend"* [P4]; *"snarky"* [P9]; and *"closed off"* [P6]. We also observed that this assessment would regularly be given in contrast to a user's expectation or willingness to engage in closer/warmer relationships. For example: *"I said 'I love you' and they didn't reciprocate"* [P2]; *"very one-sided relationship... like they didn't want to talk to me, like "here's your answer and now go away""* [P1]; *"not interested in answering back"* [P11]. Correspondingly, the "approachable" side of the scale reflect users' perceptions of an agent as emotionally warm, receptive, friendly, and willing to interact with a user. As an illustration, consider the following descriptions: *"welcoming"* [P11]; *"like I've made that familial connection with them"* [P12]; *"feels more like a friend"* [P13]; *"outgoing"* [P2]; and *"chill"* [P19].

Sentiment Toward a User

The second theme – an agent's perceived sentiment toward a user – is defined through the dichotomy of genuine vs. disingenuous agent's attitudes. While the approachability theme reflects the social and emotional closeness, the sentiment theme refers to the perceived trustworthiness in attitudes demonstrated by an agent and how focused they are on a user. Specifically, the "genuine" aspect includes descriptions of agents' actions and attitude as benevolent, caring, and altruistic, e.g. *"coming from a good place"* [P1]; *"more sincere"* [P8]; *"went above and beyond by asking if I want more"* [P3]. These descriptions also include the perceived level of care in agents' actions and motivations, e.g. an agent being *"motherly"* [P20]; *"patient"* [P17]; *"accommodating"* [P5]; *"thoughtful"* [P3]; or *"really trying to help you"* [P16]. Opposing these caring, benevolent perceptions, the "disingenuous" side of the sentiment scale reflects assessments of an agent as cunning, indifferent, or inconsiderate. For example, descriptions such as *"sly"* [P8]; *"ingenuine"* [P1]; *"don't try to figure out what I need, they think they know what I need"* [P3]; *"pretending"* [P4]; *"uses her sass as a defense mechanism"* [P7].

Figure 6.2: Number of character descriptions per agent in each category

Professionalism

The theme of professionalism, first, includes descriptions of an agent's communication style as in work settings, e.g. *"textbook assistant, no-strings attached relationship"* [P8]; *"it's weird to me when she makes a joke, because I visualize her as professional"* [P4]; *"she gives a lot of details in whatever you try to ask her, so she's a little more mature"* [P20]. Consequently, perceptions of professionalism affect a projected type of relationship with the agent: *"Siri actually feels like an assistant. Like if you had someone following you around everywhere, doing favours for you, and giving you information that you need. And Alexa, just seems like a really smart friend that you have that can answer questions"* [P14]. This theme also includes perceptions regarding an agent's professional functioning, for example, such descriptions as *"self-sufficient"* [P14]; *"really organized"* [P4]; or *"like 'I gotta get stuff done"'* [P7]. The theme of professionalism differs from the previous two themes in that it does not include a defined negative value. However, while we have

not encountered descriptions of specifically "unprofessional" behavior, we found that the comparisons between agents often reflect the differences in the degree of professionalism, especially in relation to the visualization of agents: *"Alexa has curly hair down, Google is like a bun, brown. Bun because more professional and strict"* [P10]; *"not as professional but still put together"* [P4]; *"isn't as put together, the other two are more sleek"* [P2].

Intelligence

The intelligence of an agent was mainly discussed in our data as an assessment of a positive value. For example, descriptions like *"witty with knowledge"* [P8]; *"well-read"* [P1]; or *"intelligent and sophisticated"* [P11]. It was significantly more rare for participants to describe an agent's negative intelligence (Fig.6.2), e.g. *"not intelligent"* [P5]; or *"a little on the dumber side"* [P20]. Due to the nature of the studied technology, which heavily focuses on information retrieval, the factor of intelligence might have be ambiguous notion in descriptions of an agent. In other words, a participant might be referring to the intelligence of an agent as a system or to the intelligence of an anthropomorphized persona of an agent. As an illustration, consider the following two descriptions of intelligence from the same participant: the system intelligence – *"sometimes says some weird stuff...you can tell it's trying to put sentences together, but it doesn't always make sense"* [P6] – and a persona intelligence – *"like the dumb kid that doesn't show up to class and asks everyone for help"* [P6]. In our analysis, we specifically focused on descriptions of anthropomorphized intelligence, e.g. *"a little more wise, a little more knowledgeable"* [P20]; *"just sounds really smart, like someone really knowledgeable"* [P17], and disregarded commentary on the "system" intelligence when identifying the current theme.

Individuality

The last theme of behavioral anthropomorphization is an agent's individuality. This theme refers to participants' perceptions of how much character an agent is displaying and the specifics of this character. Respectively, this theme is represented by a dichotomy of neutral versus defined individuality, where specifics of defined individuality differ for each agent.

The neutral aspect includes descriptions that explicitly highlight the ordinariness of an agent's character: *"shy...not many interests"* [P7]; *"I feel like that's an average person"* [P9]; *"just boring... less inviting"* [P11]. This factor of ordinariness would often be associated with how noticeable the presence of the agent's persona is in the environment: *"It's just some random kid that would blend in"* [P6]; *"don't particular notice them more than the average"*; [P1] *"just wanted to blend in the background and not really be noticed"* [P14].

We then analysed the descriptions of defined individuality for each agent. We found that descriptions of Google Assistant's individuality revolve around the perception of a "nerdy" persona: *"really tall and lanky, because they're supposed to be like stereotypically nerdy"* [P7]; *"warmer but nerdy... nerdy chill"* [P6]; *"awkward... lanky. I'm just picturing someone who is a little bit plain but also a little awkward"* [P5]. Among three agents, Alexa's personality was most commonly described as neutral (Fig.6.2), "blending in the background" (e.g. *"normal, nothing too out there or different"* [P7]; *"she sounds kind of dull, like she's a dull worker behind a desk"* [P11]). The defined personality of Alexa was mainly described as peculiar and rather bright, e.g. *"she has a glowing personality"* [P16]; *"should be different from the other two. I feel like maybe... a bit more eccentric"* [P13]; and *"Alexa is more of a wild card"* [P8]). The descriptions of Siri's individuality are also portraying this agent as rather unusual with an accent on being audacious and daring, which is particularly noticeable in association with the agent's perceived appearance: *"she's hard to work with... I bet she's that way because she's really hot"* [P7]; *"[with] a bit of an edge"* [P4]; *"Siri is more hipster"* [P16].

Summary

Our analysis suggests that descriptions of an agent's character and aspects of behavioral anthropomorphism fall into five categories. Each category represents a comparative scale used to assess an agent: the approachability category (emotionally/socially approachable vs. distant), sentiment toward a user (genuine and caring vs. disingenuous, cunning attitude), professionalism (degree of manifestation), intelligence (positive vs. negative), and individuality (neutral vs. defined personality). Fig. 6.2 illustrates the differences between agents' characters according to each of these five categories. In particular, Google's character tends to be described as approachable, closely followed by Alexa, but Siri was more commonly described as distant. Similarly, we observed a tendency to describe Google and Alexa as more genuine and caring, while Siri's sentiment was more commonly mentioned as disingenuous and inconsiderate. In our data, Siri's personality is commonly described as defined with particular individuality, while Alexa is noticeably more often seen as having a neutral, indistinct personality. Overall, we found it particularly interesting that the assessment of intelligence, commonly discussed in previous literature on conversational agents [416, 68, 69], is the least represented category of descriptions in our data.

6.4.2 Visualization of Agents

In the previous section we presented our results on anthropomorphized behavioral perceptions of conversational agents. In this section, we continue by discussing the visualizations of an agent's appearance. We first present the analysis of verbal descriptions of agents' physical traits[3], including physical build and age, facial expressions, hair, and clothes style (Tab.6.1). We describe the observed tendencies in the structure of descriptions in each of these categories, and discuss the specifics with regard to each agent. We then describe the analysis of the visualization exercise outcomes (Fig.6.1, Fig.6.3, and Tab.6.2) and present the consolidated results on the differences on visualizations of each conversational agent.

Physical Build and Age

In verbal descriptions of the physical traits of an agent's body, we found that some characteristics would yield rather consistent differences in perceptions of agents (e.g. height or age), while others would show no differences. Specifically, the overall build of agents would often be discussed as "average" and would not reveal any particular differences: *"fairly normal build, not super muscular... just very average"* [P1]; *"I see her as average North American built women"* [P8]; *"I picture them like thin, slim.... just average."* [P11]; *"I'm picturing a certain number of traits that are most... I don't want to say most common, but not different... very normal?"* [P13]

Figure 6.3: Collage of visualizations of agents, organized by hair style. Left to right: Alexa, Google Assistant, Siri

The factor of height, however, was a noticeable exception. In particular, participants would often note it as a comparative description, e.g. *"Same build, but Alexa maybe*

[3]Since verbal descriptions were provided in a free form, not every aspect of visualization was mentioned by each participant.

	Element	*A*	*GA*	*S*		*Element*	*A*	*GA*	*S*		*Element*	*A*	*GA*	*S*		*Element*	*A*	*GA*	*S*
Gender	Male			1	Height	Short	**2**	1	1	Age	Younger	5	6	**9**	Face	Positive Express.	3	5	3
	Female	15	14	15		Average	3	**6**	**5**		Mid-Age	4	**8**	6		Neutral Express.	1	3	2
						Tall	1	**6**	1		Older	**8**	**8**	6		Negative Express.	1	2	1
Hair Style	Short	3		**5**	Hair Color	Silver	1	1	1	Clothes Style	Casual	**18**	10	10		Glasses	4	5	
	Wavy	**6**	2	3		Blond	1	**4**	**5**		Formal	7	**16**	**11**		No Glasses	1	1	1
	Straight	2	2	2		Brunette	**9**	**8**	2		Uniform-based	2	1	3					
	Updo	4	**15**	6		Dark	**3**	2	2		Company-based	1	**4**	1					
	Down	**10**	4	6		Black	**4**	1	**5**		Individuality	4	4	**10**					
	Other	1	1	**4**		Red	1												

Table 6.1: Frequency of visual appearance elements provided through verbal descriptions of Alexa (A), Google Assistant (GA), and Siri (S).

	Element & Type	*A*	*GA*	*S*		**Element & Type**	*A*	*GA*	*S*
Hair Style	Short	1	1	**5**	Hair Colour	Silver	1	1	2
	Long Straight	4	**7**	**7**		Blonde	0	**5**	**6**
	Long Wavy	**7**	3	2		Brunette	3	**5**	2
	Updo	4	**6**	3		Dark Brown	**10**	6	3
	Midlength	4	3	3		Black	5	3	**7**
						Red	1		
Gender	Female	19	19	16	Glasses	No glasses	13	13	16
	Male	1	1	4		Sunglasses	2	2	1
						Glasses	5	5	3
Eyes	Regular	12	13	12	Mouth	Open	10	9	10
	Happy	4	3	4		Open (teeth)	5	4	4
	Sad	1	1			Closed	2	1	1
	Other	1	1	2		Neutral	3	**6**	3
						Other			**2**
Cloth. Style	V/Scoop neck	**10**	4	**8**	Cloth. Colour	Dark/Black	11	10	12
	Hoodie		**4**			Blue	3	**5**	2
	Tank top			**2**		Grey	2	1	2
	Sweater w/collar	**6**	**5**	2		White/Beige	**4**	3	
	Jacket	**7**	**9**	**9**		Other		1	**4**

Table 6.2: Frequency of visual appearance elements provided through the visualization exercise for Alexa (A), Google Assistant (GA), and Siri (S).

shorter" [P11]; or *"Siri would be the smallest, Google would be the tallest, Alexa in-between."* [P10]. The height was most commonly associated with personality and behavioral perceptions of a specific agent, e.g.: *"Alexa slightly taller because daunting."* [P8], *"maybe Google seems taller. I don't know why, but that just seems more professional"* [P13]; or *"Alexa is shorter because she's less authoritative... and, I guess, I attribute that to tallness"* [P2]. Overall, we found a tendency for the Google Assistant persona to be generally perceived with a taller average height compared to the other two agents (Tab.6.1).

The descriptions of agents' age also revealed some differences in perceptions. Similar to height, age was often provided as a comparative description, e.g. *"Alexa maybe younger, out of all of them Siri is the oldest, but not by much, in 30s or mid 20s."* [P2]; *"Siri maybe later 20's, 30's, but Alexa 10-15 years older than that"* [P12]; *"Alexa is a little older, 30 to*

40. Google is 25 to 30. I think, Siri is the youngest, 20-25." [P20]. Correspondingly, age would often be associated with behavioral perceptions of an agent (e.g. *"she is younger, fresh grad vibes. Like trying too hard"* [P6]). While we did not find any trends in a specific age group assigned to each particular agent, we observed a slight tendency for Siri to be perceived younger, and for Alexa to be perceived older (Tab.6.1): *"Alexa is older, in her late 20s. Siri and Google are in early 20s"* [P5]; *"Siri sounds younger, like late 20s. Alexa, I think, like 40s. Google sounds older... so between the two, mid 30s?"* [P9]; *"Alexa is the oldest, maybe late 30's to 40's... "* [P11].

Facial expressions

Perception of an agent's facial expressions seems, first, to be associated with how defined the agent's personality is for a particular participant, e.g. *"I guess, she looks deadpanned. If I could give her a face, I don't think I'd give her an expression"* [P4] versus *"She has an authoritative vibe... but still friendly, so still smiling, not stern"* [P2]. The specifics of facial expression often reflect the perceived approachability of an agent, for example, *"slight smile, because she is more playful"* [P3]. This is also reflected in the described differences in agents' facial expressions: *"Google just seems stressed, they've got a lot of stuff and services. Google is just focused on something, seems a little like their mind is somewhere else if you are having a conversation with them. Alexa is just maybe a little bit smiling, and Siri would be smiling a lot and has a lot of energy. In general, I can't imagine them being upset."* [P7]. Furthermore, participants commonly described agents' facial expressions based on their perceived role in an interaction: *"Alexa and Google have very neutral expressions until I prompt them, in which case it becomes a look of curiosity. In my mind, when they talk to me, it's a positive expression, and it can never be negative – I'm the client, they're doing me a service"* [P1].

Hair Style and Color

The hair styles verbally described for agents included different length, shape (*"looser curls"* [P4]; *"straight, medium/shorter length"* [P7]), style (*"tied up in a bun"* [P5]; *"down, long, wavy"* [P9]) and color (Tab.6.1). The descriptions of hair style appear to be significantly affected by perceived professionalism of an agent: *"really dressed-up person who's reporting something or like really formal, doing a really serious job. I associate that more, for females, having your hair up than down, at least for long hair."* [P17]; *"professional but also friendly, so hair in a bun or a ponytail"* [P2]; *"Alexa has curly hair down, Google has a bun, brown. Bun because more professional and strict"* [P10].

We have also encountered a number of very particular, definitive descriptions of hair, especially for Siri, e.g. *"weird short bangs"* [P15]; *"she's blonde but she always curls it. She got that Ariana Grande hair, half pony tail, really long"* [P7]; *"receding hairline"* [P6]. Such descriptions were commonly provided to highlight a perceived defined individuality of an agent. Similarly, the lack of individuality was also reflected in hair descriptions: *"hair down, long, wavy? I feel like she is an average person and this is average"* [P9]; *"a normal ponytail, I just feel like it's very normal"* [P15].

Clothes and Overall Style

We categorized the descriptions of agents' appearance style as formal, casual, uniform-based, device/company inspired, and specifically reflecting agent's personality (Tab.6.1). For example, the following description includes a formal, a personality highlighting, and a device/company inspired perception: *"Siri, I would say, formal clothes, a work outfit, like a black suit, black pants, something like that. Alexa – jeans and sneakers and a t-shirt with an effect, a sentence. Something like "females are the future". And Google would be wearing like a Google t-shirt, jeans, and sneakers"* [P19]. Similar to other appearance elements, clothing style is especially descriptive when given as a comparison between agents: *"Google would wear a straw hat, wearing like a beach outfit. I picture a summer setting, because Google is in Mountain View. I feel like Siri is more hipster, less average. Siri wears like trendy overalls. And Alexa just wears jeans and a t-shirt"* [P15].

The uniform-based category includes descriptions of clothing style based on associations with a specific profession, for example, *"I associate like flight attendant clothes"* [P9]; *"Like when you're watching TV and you turn on the news, and there's someone sitting at the anchor. I guess, someone you'd expect to be reporting... like a journalist"* [P17]; *"like a news announcer... meteorologist on the Weather network"* [P14].

The category of personality highlighting clothing includes descriptions for each agent that mainly correspond to the themes of defined individuality of each agent (sec.6.4.1). For Siri, such clothing descriptions revolve around daring personality with accentuated femininity: *"Siri would wear heels, skirt, and a crop top, like a mini skirt."* [P7]; *"I almost picture Siri in like a school-girl outfit. Like a pleated skirt... like any teenage romcom"* [P5]; *"Some really hippy, colourful dress. And a really bright dress, green-ish colour."* [P16]. Google's clothes in this category are associated mainly with the perception of a stereotypically "awkward" persona: *"I imagine to have glasses. They could be wearing jeans and t-shirt but it doesn't fit them well"* [P7]; *"Google would dress like... glasses and the weird skirts and leotards. Sweater things, sweater vest things?"* [P20]. However, Alexa's clothes in this category do not seem to be consistent. As an illustration, compare

the following descriptions: *"not regular human clothes because it doesn't feel human. Like sort of metallic clothes... maybe those reflective futuristic clothes"* [P13] and *"She's wearing mom jeans. And running shoes, because that's a very mom thing. Yeah, and glasses, that's a motherly thing too"* [P20].

Consolidation of Visualization Outcomes

We then compared the results of our analysis of verbal descriptions of agents' appearance with the outcomes of the visualization exercise. The visualizations of each agent were composed into a collage (Fig.6.3), and the number of appearance elements was calculated for each collage (Tab.6.2). We also created a cumulative visualization for each agent (Fig.6.1) to better illustrate the corresponding differences.

Our combined results suggest that Alexa tends to be perceived as being average height or slightly shorter, generally older than other agents, and wearing casual or business-casual clothes of dark or neutral colors. Her hair tends to be seen as darker, wavy, and worn down. Specifics of Alexa's appearance commonly highlight either a tendency to be unnoticeable (e.g. *"that's something I would wear if I just wanted to blend in the background and not really be noticed"* [P14]); cozy and home-associated visualization (e.g. *"a pullover sweater, skinny jeans, no shoes, I imagine her sitting on a couch"* [P4]); or an eccentric personality through futuristic clothes.

Google tends to be seen as being average height or taller, wearing either casual clothes with a focus on tech culture (e.g. hoodies), or business-formal clothes, both of dark or neutral colors. Google's hair tends to be seen of lighter color (blond, brunette) and as either long and straight, worn down or worn up (bun, ponytail), specifically associated with higher professionalism. The aspects of appearance highlighting their personality mainly revolve around the perception of a busy (*"she is not looking at me, she's busy"* [P4]), stressed (*"hair maybe gray? because of stress"* [P6]), or rather "awkward" person (*"a little bit plain but also a little awkward"* [P5]).

Siri is commonly described as being of an average height, younger than the other agents, and least often wearing glasses; wearing either casual but fashionable clothes (v-necks, tank tops, heels) or strictly business-formal style, of either dark or particularly bright colors, especially red. Siri's hair is described as short (most commonly among the three agents), or as long straight hair worn down, either blond or black, or with unique descriptions, e.g. *"black, but dyed black of white girls"* [P15]. Finally, descriptions of appearance highlighting specifics of Siri mainly portray a daring personality with accentuated femininity (e.g. *"the cheerleader trope"* [P7]).

Design	*Approachab.* app.	dist.	*Sentiment* gen.	disin.	*Individ.* neut.	def.	*Prof.*	*Total*
Information	4	2	5		2			13
Interac. Style	11	10	5	6	1	1	4	38
Voice	3	5	1	1		1	1	12
Company	2							2
Total	*37*		*18*		*5*		*5*	

Table 6.3: Associative pairs for elements of character descriptions (columns), explained through designed features (rows).

Character		*Build*	*Cloth.*	*Face*	*Hair*	*Age*	*Total*
Approachability	approach.	1	6	2		1	16
	distant	2	2	1	1		
Individuality	neutral	2	7		6		33
	defined	2	4	2	10		
Professionalism	more	3	11	2	9		29
	less		1		3		
Sentiment	genuine		2				2
	Total	10	33	7	29	1	
Design		*Build*	*Cloth.*	*Face*	*Hair*	*Age*	*Total*
	company	2	6		1		9
	int. style		3			4	7
	voice		2			4	6
	Total	2	11		1	8	

Table 6.4: Associative pairs for elements of visual appearance (columns), explained through character and designed features (rows).

6.4.3 Associations between Character and Appearance

Earlier, we discussed the structure and specifics of anthropomorphized perceptions of agents' character, as well as visualizations of their physical appearance. To further operationalize these results we now explore the associative relationships between different descriptive elements. For the analysis of these associations, we identified self-reported explanations (n=167) provided by our participants within their anthropomorphized descriptions (e.g. *"hair is down and wavy because she isn't as put together"* [P2]). In particular, we wanted to find the associations between behavioral and visual perceptions and the designed (objective and intended) features of each agent.

We first explored the designed features mentioned by participants as associative explanations, and identified four corresponding categories. First, (1) the category of provided information, including the amount of information and level of details (*"they give more info and so it seems like they care about you understanding what they're saying"* [P2]), as well as the very fact of providing the information (e.g. *"Siri is just cold, she copped out on a lot of answers and that made it more distant"* [P8]). Second, the designed (2) interaction style of an agent, such as responsiveness to prompts (e.g. *"I always have to be very clear with what I want from her, so she doesn't feel as flexible as a person"* [P1]), prompting users back (*"seems closed off, because doesn't do any prompts back like 'how are you' and stuff"* [P6]), and the tendency to use jokes (*"when I asked what's her favourite colour, she told me a joke. So yes, way more playful"* [P3]). The third category of features refers to agent's (3) voice characteristics, e.g. *"More enthusiastic? like there's more inflection in her voice, sounds a bit more fun"* [P13]. Finally, we included the category of associations motivated by the (4) perception of the company, e.g. *"Alexa is more chill. Maybe because it's associated with a brand that reminds me of books and stuff"* [P19].

We then separately considered the associative explanations for elements of character (Tab.6.3) and visual appearance perceptions (Tab.6.4). We found that different categories of character perception are self-reportedly motivated only by designed features of agent technology (Tab.6.3), while visual appearance is motivated by both designed features and – predominantly – by character perceptions, with the exception of the agent's age. Note that the factor of intelligence did not appear in any of the elicited associative explanations.

In the character perception, the designed features are mainly affecting (Tab.6.3) the perception of agent's approachability, followed by the assessment of their sentiment toward the user. Both character categories are predominantly influenced by the interaction style of an agent, and the information they provide. Further, agent's approachability is associated with the agent's voice characteristics, e.g. *"less approachable, because the voice is just dull, doesn't seem interested in answering back"* [P11]. Overall, among the identified designed features, the agent's interaction style is most commonly associated with character perceptions, including – in order of extent of influence – agent's approachability (n=21), sentiment (n=11), professionalism (n=4), and individuality (n=2).

The visual appearance was predominantly motivated by the character perceptions (n=80). The most pronounced influence on visualization comes from the perception of the agent's individuality, both as defined (n=18) or neutral (n=15). The next most influential factor is the agent's perceived professionalism (n=29), followed by the agent's approachability, both when assessed as approachable (n=10) and distant (n=6). Interestingly, the perceived sentiment toward a user does not appear often in association with visual appearance, and intelligence of an agent does not appear at all. The most common

aspects of appearance motivated by an agent's character include clothing style (n=33), followed by the agent's hair (n=29), and physical build (n=10). Perceived age of an agent, however, is mainly motivated by designed features, specifically by the interaction style and voice features of an agent. Designed features also seem to noticeably affect visualizations of an agent's clothing style, especially associations with the corresponding company. The design features mentioned as affecting visualizations include interaction style, voice, and the agent's parent company, while the information factor did not appear in our data.

6.5 Discussion

Our findings suggest that the purposefully anthropomorphized behavioral and visualized perceptions of conversational agents yield structural consistency, discriminating differences between agents, and are linked with each other and system features though underlying associative relationships.

We found that the descriptions of behavioral anthropomorphism of conversational agents can be structured around five categories – approachability, sentiment, professionalism, intelligence, and individuality – each represented by a corresponding dichotomy. These results are particularly revealing when considered in the context of related theories. Exploring the concept of humanness and dehumanization, Harslam [198] describes two distinct senses of humanness: human uniqueness, believed to be acquired and to vary between people, and human nature, seen as fundamental characteristics "embedded" in the person, including cognitive flexibility, emotional responsiveness, individuality, depth, and interpersonal warmth. Harslam argues that the combination of attributed opposite characteristics (inertness, coldness, rigidity, fungibility, and lack of agency) represents a mechanistic dehumanization – a view of others as object-like or automaton-like. First, clear parallels between Harslam's human nature characteristics and some categories suggested in this project (e.g. approachability – interpersonal warmth, sentiment – emotional responsiveness, individuality) validate the identified categories and the authenticity of the anthropomorphized nature of our data. On the other hand, the differences in the categories highlight the specifics of behavioral anthropomorphism of conversational agents. For instance, the behavioral category of professionalism suggests the tendency to project the role-based relationship onto the agent. While previous research also separates users' assessment of agents' access to knowledge and intellect [416], our results show that the anthropomorphized intelligence of an agent is associated with knowledgeability, rather than cognitive flexibility, but the complexity of system performance affects either perceived professionalism or agents' sentiment.

The identified categories of behavioral descriptions also allowed us to compare the anthropomorphized perceptions of agents with respect to each category. For example, Siri's sentiment was predominantly described as disingenuous and cunning, while Alexa was described as genuine and caring. These particular results support findings from an earlier study [53] that assessed emotional responses to Alexa, Siri, and Cortana by measuring participants' biometric data, and concluded that Alexa generated the greatest emotional connection with participants. Contrarily to the sentiment perception, Alexa's individuality was commonly described as neutral and ordinary, while Google's and, especially, Siri's are considered to be more defined and pronounced. Finally, the variations in the relative "distance" between agents in each behavioral category both support the legitimacy of thematic clustering which defined these categories, and provide material for further exploration or the mechanisms underlying the differences.

Our results on the visualizations of purposefully anthropomorphized perceptions of agents demonstrate rather consistent differences in participant perceptions of agents' appearance, particularly in factors of their height, age, hair, and clothing style. We first would like to acknowledge that both character, and especially visual perceptions, elicited from the study, demonstrate an extreme tendency for displaying biases and social stereotypes. While this tendency is well established in anthropomorphized perceptions of computational systems [360, 59, 60, 147, 325], it might appear particularly defined for conversational agents due to their integration into users' personal and social environments [478]. Furthermore, a recent policy paper, produced by UNESCO [498], suggests that conversational agents potentially promote gender stereotypes by acting submissively in response to a user's displayed gender bias or even abuse. We have intentionally skirted analytical lenses in this project, e.g. feminist theory, that might probe these aspects in more detail; this is an obvious avenue of future work.

In this project, we rely on a comparative analysis, and present the specifics of each agent's visualization with the main purpose of demonstrating that visualization differences are present and consistent. Presenting these results, we mention behavioral connotations observed for particular forms of visual elements, e.g. hair worn up as a reflection of higher professionalism. However, due to the nature of the study, our data does not provide enough evidence to reliably establish detailed correlations. In essence, while our findings describe the structure of character and visual anthropomorphized perceptions, and demonstrate that they are interdependent, further research is required to confirm and expand on the understanding of specific dependencies between aspects of the agent's character perception and particular shape of visual elements. If established, these correlations could open exciting opportunities both for the system's feature design (e.g. designed answers referring to agent's visual appearance), as well as for the development of new projective investigation

tools [268, 155] for research on perceptions of virtual agents.

As a step toward establishing these dependencies, we examined the self-reported associations between different categories (but not detailed forms) of anthropomorphized perception, and their relation to system features. Our results suggest that categories of character perception are only associated with the designed features of agents (information format, interaction style, voice characteristics, and agent's company), while visual appearance is motivated by some designed features (interaction style, voice characteristics, and agent's company), but mainly reflects perceptions of an agent's character. This both suggests the interdependence of agent character and visual conceptions, and the dominance of the behavioral anthropomorphized perception of conversational agents. These results also illustrate the differences of the relevant impact of system features on anthropomorphized perceptions, affecting mainly perceptions of agent approachability, sentiment, and age. Visual perceptions are mainly associated with the agent's perceived individuality, professionalism, and approachability, and rarely with the agent's sentiment. Understanding these relationships between the identified categories of features, behavioral perceptions, and visualizations opens new avenues for a guided design of conversational agents.

6.5.1 Limitations

First, we acknowledge the effects of the diverse levels of participants' experience with conversational agents. Thus, while we found structural consistency in the descriptions of personality and visual appearance, we focus on reporting the specific attributes associated with a particular agent predominantly as an illustration of a theme. Correspondingly, our study design included the interaction session. While the main goal of the session was to equalize the recency of relevant stimuli for comparison rather than attempting to equalize the overall experience, the length of interaction sessions might also introduce some limitations in the results. Finally, we acknowledge that the benefits of standardizing the medium for visualizations for reliable comparisons can introduce limitations in the freedom of expression. To address the potential priming effect of the medium, we first conducted interviews, which, as reported, frequently showed *a priori* detailed perceptions of agents.

6.6 Conclusion

In this chapter, we discussed the results of an exploratory study on the structure and interconnectedness of the anthropomorphized behavioural perceptions and visual conceptions of conversational agents. While, to the best of our knowledge, this study is the

first attempt to identify patterns in the purposefully anthropomorphized perceptions of conversational agent technology, the consistent differences revealed through a comparative analysis of three agents suggest the importance of this research direction for the design of conversational agents.

The primary goal of this chapter was to understand the human-like features of these agents – appearance, personality and intelligence, for example. However, another aspect of interaction with these agents is what we expect them to be able to do, the roles we want them to play, and our concerns about their actions when they serve in these roles. We probe these questions in the next chapter.

Chapter 7

Project 6: Conceptual Model of the Conversational Agent Work Process

In our final project on communicating with technology, in this chapter we look at the role users want conversational agents to play. Alongside this, there are obvious privacy and trust issues surrounding use of these agents, and understanding how users perceive both the role of these agents and the risks of their use sheds further light on our understanding of what role these conversational agents can play in support of their users.

7.1 Introduction

Voice assistants, also known as conversational agents (CA), voice activated virtual assistants (VAVAs), or intelligent personal assistants (IPAs), have prompted a surge of interest in research on interaction design [82, 305, 289, 272] and on adoption [285, 416, 233, 364, 39, 70] of this technology. One particular area of research in CAs is in the types of tasks that people perform through CA systems.

Recent rapid consumer adoption of smart speakers [482, 483, 352] reveals a non-direct relation between the tasks users perform with these agents and the set of virtual assistant capabilities, commonly called "skills" or "actions" [500, 353, 416]. Both academic [285, 416, 39] and non-academic [482, 271, 342, 78, 70, 69] sources demonstrate that voice assistants are being used for a variety of selectively assigned daily tasks including playing music, answering general questions, setting alarms and timers, placing online shopping orders, and controlling networked devices. However, while preferred tasks might change over time

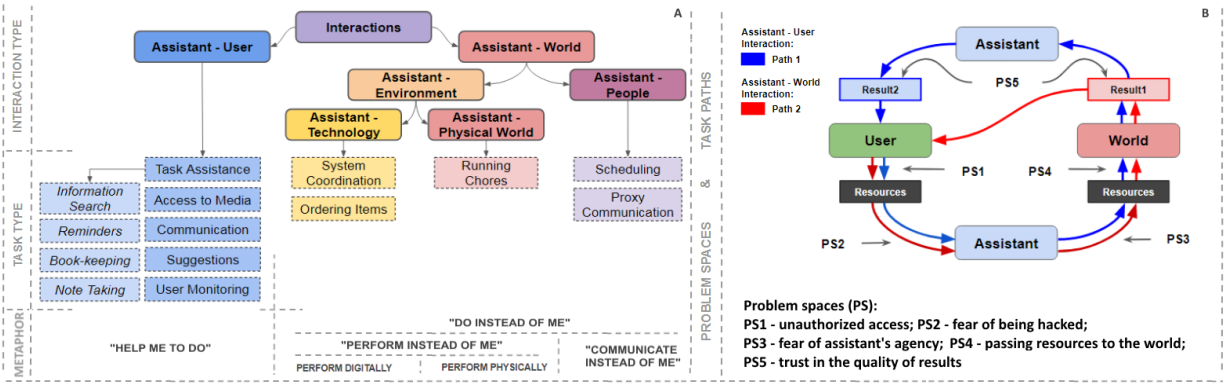

Figure 7.1: Conceptual model of the conversational agent work process. Left - classification of interaction types and tasks. Right - task execution paths and corresponding problem spaces (PS).

[39], multi-country exploratory research suggests that the majority of voice assistant users only tend to carry out between one to three different tasks or use cases [353, 97]. Yet, CA technology is designed to support a significantly greater variety of tasks than are currently being observed in common adoption – in addition to over 80,000 pre-designed skills/actions available for popular voice assistants [352], companies also provide users and third-party developers with an option of creating their own actions for the assistants [23, **?**, 500]. This increase in available tasks coupled with the limited use cases employed by users [353] further exacerbates the imbalance between observed and potential CA actions.

One issue leading to discrepancies between the accessible and adopted capabilities of CA technology is the challenge associated with low discoverability of agents' skills and actions [500, 416]. Additionally, adoption patterns seem to be mediated by an underdeveloped mental model of CA systems [272]. Furthermore, while users commonly engage in testing the system's capabilities [416, 289], they also tend to assume that assistants have a low level of competence for complex tasks [68]. The significance of understanding users' expectations is even greater, as research demonstrates that users' uncertainty about system features and discrepancies between expectations and the reality of the system's operation lead to frustrating experiences, lowered expectations, and non-use [289, 97]. Finally, security and privacy vulnerabilities of CA systems [523, 99, 166, 218, 528] find reflection in corresponding users' concerns [264, 138, 68, 297]. Specifically in the context of CA adoption as part of intelligent home systems, these concerns fit into a broader system of privacy beliefs, such as fear of surveillance and loss of control over personal data collection [31].

Unsurprisingly, these factors are interdependent, and thus affect the adoption patters holistically and systematically. For instance, Lau et al. [264] explored people's reasoning

around adoption of smart speakers, as well as related privacy perceptions and privacy-seeking behaviors. They showed that active users have few privacy concerns, but also lack an appropriate understanding of the system and rely on "the socio-technical context in which smart speakers reside", while non-users tend not to trust or see the utility of smart speakers. Similarly, there is evidence that CA adoption is not a question of knowledge about the service, as most non-users are aware of it, but might be more related to a "missing trigger from knowledge to actual use" [353].

Taken together, the above findings demonstrate that, although there exists a significant diversity of available skills for agents, users actively engage only with a small subset of these skills, and the use of smart speakers tend to decrease or disappear completely with time. In addition, both users and non-users experience the lack of understanding of the system's functioning and capabilities, which leads to fears and concerns, frustration, and other unfavorable predispositions. In combination, these adoption characteristics might be considered as a symptom of a novel technology being introduced to the user without an appropriate exploration and understanding of the corresponding design space. Indeed, the overview of the related literature demonstrates that researchers have mainly explored domain-specific design spaces for this technologies, such as pedagogy (e.g. [463, 226]), healthcare (see [262] for the review), care-giving (e.g. [383, 513]), cooking (e.g. [486]), work-related activities (e.g. [505, 233, 236]), etc. However, while CAs such as Siri, Alexa, or Google Assistant are considered to be general-purpose, as opposed to domain-specific assistants [178], there is a noticeable paucity of exploration of the design space in which these general-purpose agents should be embedded. In other words, we currently have poor understanding of what users' general needs should be addressed through CAs. As a result, the CA systems commonly adopted in the consumer market suffer from low functional dependency [97], and consequently, tend to be perceived as "useless" [97, 264], leading to the alarming retention rate of voice assistant technology [482, 483].

Thus, we argue that there is a pressing need for a systematic and structured view of the tasks that should be supported through CA. This need is further motivated by the evidence that the set of tasks associated with agents is not uniform, and users' interaction experiences vary depending on which tasks they perform using CAs [233, 83]. While current research on CA adoption provides descriptive lists of common tasks [416, 285, 70], and acknowledges the effect of users' beliefs on this adoption [138, 289, 68], it appears that no comprehensive classification of user-envisioned CA tasks has yet been presented [353].

However, providing a comprehensive classification, arguably, only addresses a subset of the challenges associated with CA adoption; in fact, we suggest that, alongside task classification, CA skills should be considered in light of users' functional needs and their mental model, expectations, and concerns regarding CA technology. The importance of

consideration of agent's tasks in relation to users' perceptions is highlighted by the fact that user expectations of voice agents remain far from the practical realities [289], and their trust in CA systems is built on incomplete understandings of task performance processes and associated risks [272, 264]. Moreover, to address the issues of discoverability [500, 416], expectations-matching [416, 289, 68], and non- or declining use [353, 289, 97], we need to consider not only the performance beliefs of current users, but the general audience's conceptualization [463] of the possible tasks, task performance, and associated pitfalls.

In this project, we capture and systematize user envisionment of the CA's possible tasks with performance expectations. Reeves [382] defines envisionment as a "broadly future-oriented aspect of technology design which mixes fictions, forecasts, extrapolations or projections into societal visions for technological progress" [382]. Traditionally, envisionment is understood as a method which typically involves design specialists; however, some researchers have also utilized envisionment through a participatory scenario-based approach as co-design explorations [392, 473]. In this research, we use "user-envisioned tasks" to refer to the tasks elicited from our participants through a probe-based envisionment exercise in the form of a thought experiment built around a hypothetical scenario. This methodology allowed us to benefit from a user-centred design approach [473, 61], while exploring participants' needs and perceptions beyond their direct system familiarity [216]. We present our findings as a conceptual model of the conversational agent work process (fig.7.1), which includes four components:

- First, a classification of interaction types, based on a combination of users' involvement in the primary activity and the nature of interacting entities.
- Second, a corresponding typology[1] of potential CA tasks. The task categories were formed based on the interaction of four parameters: content, required device infrastructure, required system intelligence, and the temporal aspect of the task.
- The third component is a schematic representation of two task execution flows associated with the suggested interaction types. These flows reflect two major scenarios from the perspective of resource dynamics and task outcomes, and allow us to conceptually bridge the interaction types and users' concerns.
- Finally, the fourth component of the model is a classification of users' concerns mapped onto the task execution flows based on the nature of concerns and their correspondence to system architecture and design components.

The conceptual model offers a systematic understanding of CA tasks in association with the user's perceptions and beliefs about the system. The implications of this struc-

[1]Here typology is used to refer to multi-attribute classification on a nominal scale.

tured synthesis contribute, first, to the understanding of how to develop CA skills and to appropriately form the skills templates. Furthermore, the holistic view on the expectations on tasks and task performance allows the design of relevant system feedback and appropriate representation of system intelligence [289]. Additionally, it informs directions for further research on leveraging contextual and personal cues to proactively identify and recommend skills to the users at the correct moment [500]. Finally, the consideration of the users' expectations on task performance in relation to the associated pitfalls informs both how to support the development of a suitable mental model and to better address users' trepidation regarding these agents.

In the remainder of this chapter, we first describe a study of CA envisioned tasks and related processes that we leverage to generate the model. Our model synthesis includes a joint discussion of the classification of interaction types and the typology of tasks, task execution flows, and clusters of user concerns. We then discuss the implications of the model for the research and design of conversation agent systems.

7.2 Related Work

Research on the use of intelligent assistants shows that users tend to selectively assign certain types of tasks to this technology [285, 416, 39]. Exploring ethnographic aspects of adoption of CAs, Sciuto et al. [416] examined how households integrate Alexa into their lives. Their results demonstrated that, for in-home participants, Alexa was naturally integrated into daily routines, e.g. feeding children, getting dressed, showering, playing musical instruments, cooking, and getting ready for bed. Collecting data through an online questionnaire and diary methods, Lopatovska et al. [285] found that across all age groups, Alexa was primarily used for checking weather forecasts, playing music, and controlling other devices. Authors note that several participants reported using Apple Siri and Google Now applications in addition to Alexa for similar purposes, except for controlling other devices. Similarly, the results of large-scale user surveys [70, 353, 232] show that CAs are most commonly used for simple information retrieval, checking the weather, and interpersonal communication.

Beyond listing tasks and activities commonly performed with CAs, there is a noticeable paucity of attempts to organize these tasks into structured typologies. Among few examples, Catrambone et al. [83] proposed a framework for anthropomorphic interface agents, emphasizing features of the agent, the user, and the task the user is performing. The framework suggests that tasks performed with an agent's help might be classified along some or all of the following dimensions: objectiveness (opinion-based recommendations

versus facts), intent (learning goal versus familiar steps), and other task-related variables, such as time pressure, duration, and consequences of the quality of task performance. Another example of categorizing agent's tasks comes from non-academic research and revolves around the complexity of an activity [70], grouping CA tasks into four categories based on the number of required steps: simple actions which require one step to complete (e.g. setting a timer), multi-step tasks which require several stages to complete a process (e.g. placing an online order), multitask activities involving several activities and applications (e.g. creating a list of phone numbers for people whose emails you have not read), and research activities which require synthesizing and analyzing multiple information sources (e.g. finding the best hotel based on a set of criteria). However, while these frameworks provide a useful way of categorizing agents' tasks, they are not sufficient to describe the design space for CAs and to address the question of users' corresponding functional needs.

Interestingly, in everyday adoption, CA users seem to continue using basic tasks [97, 353] despite growing technological capabilities of the system [352], and the ability for developers and users to create custom skills for specific tasks [23, **?**]. Thus, we observe the users' low functional dependency on CAs [97], the worrisome retention rate of voice assistant technology [482, 483], and an overall low interest in using the system [97, 264, 353].

One of the roots for the discrepancy between available and adopted skills seems to be users' lack of confidence in a CA's performance – users tend to assume that assistants have a low level of competence for complex tasks and find interactions with them to be socially awkward [68, 69]. Indeed, despite being effective for some complex activities [463], in other scenarios intelligent agents are shown to fail to appropriately support complex interactions. For example, targeted evaluations in the domain of health demonstrated that when asked simple questions about mental or physical health, and interpersonal violence, responses of popular CAs were inconsistent and incomplete [310]. Reis et al. [383] evaluated functionalities of several voice assistants (Amazon, Google, Microsoft, and Apple) and their compliance with the interaction scenarios for adoption by the elderly. The interaction scenarios included basic greeting, email management activity, social and family events management (e.g. notifying the user about a social event), and social games (e.g., cards, trivia). Their results demonstrated uneven possibilities for the tested CAs to fulfil the majority of interaction and activity objectives. Consequently, users engage in testing agents' competence. Sciuto et al. [416] showed that in-home participants tested Alexa's capabilities by asking questions from three categories: access to knowledge (factual questions), intellect (sophistication of commands), and personality (naturalness of responses).

Another factor affecting CA adoption is an underdeveloped mental model of the system [68]. Luger and Sellen [289] explored usage patterns, motivations, and expectations on the everyday use of IPA. They showed that users' uncertainty about system features and

abilities leads to frustrating experiences and decreased usage. Authors suggest that, to address the effects of mental model shortcomings, IPA system feedback should be improved, and the design should be more transparent regarding system intelligence [289]. Further highlighting the role of users' mental model, Li et al. [272] identify three common sources of failures in human-agent conversations: unknown concepts, out-of-domain tasks, and wrong fulfillment means or level of generalization in task execution.

Research also demonstrates the complex connection between user satisfaction, perception of a CA, and the types of performed tasks. For instance, users' perception of an agent's personality, helpfulness, and intelligence seem to be significantly influenced by the task performed with an agent (carrying out procedures vs. providing opinions) [83]. Kiseleva et al. [233] explored user perception of Cortana performing tasks of device control, web search, and structured search dialogue, and found that the concept of user satisfaction varies across tasks [233], and for some tasks depends on task completion, for others on the amount of effort spent on a task. Moreover, the interaction experience might potentially be more important to users' satisfaction than the interaction output. For instance, Lopatovska et al. [285] showed that users reported being satisfied with Alexa even when it did not produce sought information. Similarly, users' satisfaction and relationship with conversational agents are shown to be significantly mediated by perceived trust and enjoyment during interactions [267]. Additionally, an interaction experience with an intelligent agent might be influenced by contextual aspects of interaction. For example, Bickmore and Cassel [46] found that participants' perceptions of the agent are sensitive, among other factors, to the agent's interaction style (social vs. task-only dialogue).

Finally, the adoption of CA is affected by users' privacy concerns; i.e. issues with the systems' vulnerability to security and privacy attacks [99, 166, 523] find reflection in CA usage patterns [297, 402]. For instance, users' privacy concerns include the belief that the CAs are always listening to users, that the information is recorded and transmitted to the cloud in full (rather than abstracted) form for further interpretation, and discomfort with using agents when an error or a misunderstanding could have noticeable consequences [68, 297]. The exploration of social issues around the use of CA demonstrated [138] that people tend to avoid using voice input in public settings due to privacy concerns. Similarly, research on social aspects of smart homes adoption [31] shows that although users acknowledge benefits of having this technology, there is a fear of loss of control, and perception of such technologies as "big brother-like".

Thus, previous research on the adoption of CA has predominantly focused on presenting tasks commonly performed with agents, and explored a number of parameters affecting the adoption of this technology, including confidence in agents' abilities, users' understanding of the system's working processes, satisfaction from interaction experiences, and privacy

concerns. Research has also discussed the challenge associated with low discoverability of agents' skills and actions [500, 416]. While perusing a list of tens-of-thousands of skills does little to foster a better understanding of CA systems, White [500] suggested leveraging context to proactively recommend new skills. However, he also acknowledges the challenges of insufficient context to infer, lack of trust to give CAs access to contextual information, and intrusiveness of recommendations, all of which may limit skills recommendation.

To summarize, addressing challenges of low discoverability, user's poor understanding of the system work process, and mismatches between users' expectations and reality is predicated upon developing a systematic view of the user's understanding of appropriate CA tasks in the context of users' functional needs, expectations, beliefs, and concerns regarding CA technology. In the remainder of this chapter, we present a four-component conceptual model of the user-envisioned work process of conversational agents to expand on our understanding of functional types of conversational agent's tasks, and how they relate to the perceptions of the task execution process and users' system of beliefs.

7.3 Study Design

7.3.1 Study Description

To recap, our interest in applying a holistic approach to the CA design space exploration is motivated by the fact that users have low functional dependency on CAs [97] and tend to actively carry out only a small number of available CA tasks [353, 285, 416, 39]. Previous research suggests that this issue might be associated with the low discoverability of agents' skills for current users [500, 416], an underdeveloped understanding of CA systems for both users and non-users [272, 289, 68, 264], and related concerns about system performance [264, 68, 297]. Correspondingly, the discrepancies between users' expectations and the reality of the system's operation cause a significant decrease in CA use or, even, non-use of CAs [289, 97, 483]. In the introduction, we noted that, at a higher level, the problem appears to stem from the fact that there is a troublesome lack of exploration of the design space for CAs that considers users' functional needs in the context of their mental models, beliefs and expectations.

To address this gap in systematic understanding of user perceptions of CA tasks and overall work processes we designed a qualitative study to collect the data set of envisioned tasks associated with CA systems, as well as related expectations and beliefs regarding how system functions. We conducted a "thought experiment" in the form of an extensive,

two-phase, open-ended survey. The survey began with a scenario description, followed by four open-ended questions. In the opening hypothetical scenario, we asked participants to imagine that they had a conversational agent in their home, where the agent did not have any technological limits to its functionality. Our motivation for the "thought experiment" format was three-fold: to explore the needs beyond "known" skills, to use a scenario-based approach, and to include both current users and non-users of CA systems.

First, shifting the focus from users' existing practices, hypothetical creative scenarios in thought experiments allow an unbounding of participants' thoughts, needs and perceptions from specific contexts and system familiarity [216]. Thus, for the purpose of this study, the chosen method provided an access to envisioned use cases as opposed to gathering purely known skills, because the understanding of known skills is often incomplete [500, 353, 416].

Second, one of the strengths of hypothetical creative scenarios in HCI research is in how it facilitates user-centred design through probing participants' scenario-based envisionment [473, 61]. Thus, the format of a thought experiment allows us to gain a more holistic understanding of participants' perceptions of CA tasks.

Finally, the hypothetical nature of the scenario opens an opportunity to work both with current users and non-users of CA technology. Building on previous research, we were particularly interested in collecting information from both users and non-users because there is evidence that, for users, few skills are actively used [353, 416], CAs are often used less or not at all with time [482, 97, 289], expectations and realistic system performance are misaligned [416, 68], discoverability issues persist [500]. At the same time, non-use seems to be related to a "missing trigger to actual use", rather than the lack of awareness about system capabilities [353]; however, non-use might also be associated with an underdeveloped mental model of CA technology [264, 289].

We wanted to identify the types of tasks or functionalities that participants find relevant for our described scenario. We also wanted to explore perceptions of tasks in relation to the functionalities that might impact users' trust attitudes towards these agents. Thus, following the presentation of the scenario, we first asked participants what would they feel (1) comfortable and (2) uncomfortable with a CA being able to do, and why. We intentionally formulated these questions both generally, and complementarily as positively and negatively phrased queries. The goal of general phrasing of the questions was to provoke a creative approach, while keeping the user-centered perspective and connotative focus on participants' needs as opposed to a pure focus on their knowledge of already existing skills. The combination of positively and negatively phrased questions allowed for prompting comparative and elaborative answers, providing us with a more nuanced and contextualized understanding of participants' expectations, system of beliefs, and associated levels

of tolerance. Finally, building on the previous findings that perceptions of an agent's personality, helpfulness, and intelligence are influenced by the type task performed [83], the questions asked about participants' comfort, instead of what they would "want" or "need".

We continued by asking participants what behavior or traits of an assistant would (3) positively and (4) negatively affect their trust in this technology. Similar to the first two questions, the complementary positive and negative phrasing provides additional elaboration and control for the data collected through questions 1 and 2. We formulated the questions around participants' trust to capture the potential diversity of perspectives in their expectations and overall system of beliefs. Specifically, without leading participants through the phrasing of the questions, we wanted to prompt their privacy expectations [297, 68, 264], aspects of their trust in the agent's performance and intelligence [416, 289, 83], and expectations related to the anthropomorphized perceptions of agents [83, 363, 254]. For instance, there is evidence that the magnitude of anthropomorphization of agents and their social behavior shapes users' trust in the system, but anthropomorphic features can also increase trust resistance [116]. Furthermore, research suggests that users' satisfaction and relationship with intelligent agents are mediated by perceived trust and enjoyment during interactions [267], which are sensitive, among other factors, to the agent's interaction style [46]. Respectively, the interaction experiences with conversational agents might potentially be more important to users' satisfaction than the actual interaction output [285].

In addition, we collected information about the level of experience that our participants had with the use of CAs. We asked "Do you or have you ever used similar voice assistants" (multiple choice question, Tab. 7.1), and (5), "If not, is there any particular rationale for that?"[2] We concluded with an open-ended question on (6) anything else a participant would like to share with us on the topic of the survey (intentionally vague to capture any additional details participants might wish to share on agents in general, the scenario, or their personal experiences).

7.3.2 Study Procedure and Participants

The data collection was performed using an online survey platform in two waves (n=111 and n=161) with nine months between waves. Our decision to run the second wave of data collection was motivated by the rapid growth of consumer adoption of smart speakers [482, 352]. For instance, while the first smart speakers were launched several years ago,

[2]While the question asks for rationale for not using voice assistants, some participants who had used voice assistants also answered this question, and we included all answers in our data set.

recent adoption reports demonstrate the steady expansion of the user base and the number of smart speakers per household, accompanied by a decline in daily and monthly active use, potentially reflecting the spread of the ownership base beyond early adopters [482, 483].

We acknowledge that leveraging the full advantage of scenario-based user research requires a complete participatory co-design process, including the back and forth interaction between participants and designers. However, since our interest was in gathering an understanding of appropriate CA tasks in the context of participants' functional needs, expectations, beliefs, and concerns regarding CA technology, in this particular study we focused on collecting data through probe-based user envisionment, rather than a full co-design process. Correspondingly, we chose to run the study in the form of an open-ended online survey, for two reasons. First, previous research has demonstrated the effectiveness of collecting design inputs through open-response questions from large audiences online (e.g. [183]). Second, crowdsoursing the scenario-based probes through an online platform allowed us to access a larger and more diverse audience, thus forming a more representative sample of current and potential users [482, 483].

In total, we collected responses from 272 participants (see tab.7.1 for demographics). Focusing on the envisionment of the CA's possible tasks and role in users' lives, we wanted to avoid constraining the study to the current state of technology. Thus, while we control for the level of experience using CAs, we intentionally consider participants with diverse levels of experience for three reasons. First, it allows us to broadly explore a design space for CA technology by capturing perceptions of potential tasks and their execution process as opposed to already commonly adopted tasks. Second, it reveals the perceptions of these tasks in relation to users' beliefs and concerns, covering the cases on non-use caused by these beliefs [264]. Finally, previous findings suggest that there is more than two times the population who knows about CAs but does not use them than those who know and use them [353]. Correspondingly, the wide awareness of CA technology and the popularity of it in media coverage suggest that even non-users might develop initial mental models which are challenging to control for. Thus, all participants were familiar with the CA technology, but had diverse levels of experience and activity of use.

7.3.3 Data Preparation

Through both data collection waves we collected 1030 meaningful[3] qualitative responses for the four core open-ended questions, Qs 1 - 4 (262, 259, 256, and 253 responses, cor-

[3]Due to the nature of the data collection process, some participants skipped or provided meaningless responses to some questions.

	Gender		Age		Use of voice assistants	
W1 (n=111)	Female	56.19%	18-29 y.o.	24.76%	Never	43.24%
			30-44 y.o.	28.57%	A couple of times	32.43%
	Male	43.81%	45-60 y.o.	24.76%	Often	18.02%
			¿60 y.o.	21.90%	All of the time	6.31%
	Skipped	5.41%	Skipped	5.41%	Skipped	0
W2 (n=161)	Female	53.40%	18-29 y.o.	14.29%	Never	42.86%
			30-44 y.o.	19.88%	A couple of times	33.54%
	Male	42.86%	45-60 y.o.	27.95%	Often	16.77%
			¿60 y.o.	34.16%	All of the time	6.83%
	Skipped	3.73%	Skipped	3.73%	Skipped	0

Table 7.1: Participants: the first (W1) and second (W2) waves of the study

respondingly). Additionally, we collected 170 meaningful responses on experience using voice agents (question 5), and 64 meaningful responses for the last question on a participant's additional thoughts on the survey topic (question 6). Thus, in total, we analyzed 1264 open-ended responses from 272 participants. We started by separating data into two bodies: examples of tasks (e.g. "ordering groceries" [P114], "drive my car" [P163], "Look up directions" [P253]), and behavior descriptions (e.g. "listening to conversations" [P81], "Asking me very personal questions" [P104], "Predicting my behavior and storing my information" [P218]) and expectations (e.g. "I have problem trusting things made by people as they are not perfect" [P56], "because it is not a person and it might be hacked" [P4], "because it could be used against me" [P257]).

Within the task example responses, often, a single response would include multiple task examples. Thus, in preparation for the analysis, we distinguished each task example and analyzed them separately. For example, an answer *"turn on lights, control temperature, turn on and control TV" [P52]* would be considered as three task examples. Similarly, answers would often contain a combination of types of information. As an illustration, consider the following response for the second question of the thought experiment: *"Ordering something, looking something up online, making appointments, paying bills. These are complicated and complex actions that could be impossible to get 100% correct in every case. They also require my analysis, judgment, intervention, decision making" [P93].* In preparation of our data, we would consider this response as including four task examples and two rationales. Each element of data was indexed in correspondence to a participant, a question, and a sentiment (comfortable/positive affect vs. uncomfortable/negative affect).

In total, the first body of data included 665 task examples, and the second body of data included 586 behavior and expectation descriptions.

7.3.4 Data Analysis

We first completed a full analysis of data from the first wave of collection. Following a grounded theory approach [107], we coded data in each body of responses in two phases. First, we performed incremental open coding. Open coding was performed by the primary author, and codes were discussed and refactored in consultation with additional members of the research team. Next, we performed axial coding, seeking relationships between the lower-level coding. We then conducted the second wave of the survey using the same scenario, questions, and data collection platform. We prepared data for analysis the same way as for the first wave. Seeking validation for the coding scheme developed earlier, we applied it to the new data, comparing degrees of descriptiveness, thematic precision, and appropriateness of generalization level between both data corpora. This procedure allowed us to confirm our initial coding scheme and, thus, we proceeded with analysis using combined data from both information collection waves.

7.4 Interaction and Task Classifications

In this section, we present the first two components of the conceptual model of CA work process: a classification of six interaction types, and a typology of task activities, which includes ten categories and four subcategories. We discuss these two components jointly since they both were formed based on the analysis task examples data. The top level of Fig.7.2 represents the interaction types, and their corresponding interaction metaphors are pictured at the bottom level. The middle level of Fig.7.2 shows categories of task activities in correspondence to each of the identified interaction types.

To generate the relationship between task categories and interaction types in Fig.7.2, we analyzed the first body of responses, which contained participants' descriptions of tasks associated with CAs. Using incremental open coding, we identified ten task categories (fig.7.2, middle), formed based on the interaction of four parameters: (1) content, (2) involved device infrastructure, (3) level of required intelligence, and (4) the temporal aspect of the task performance. For example, the task category "reminders" requires access to the user's calendar, while "ordering items" requires access to user's financial resources and to retail sources. The required intelligence level refers to the involvement of the secondary processing of the information by the agent (e.g. *"Summarize news" [P250]*). The temporal aspect reflects the immediacy of the appropriate outcome. For example, the task category "book-keeping" implies prolonged task performance, while "access to media" implies immediate results. While the majority of task examples were provided

in the context of a positive sentiment (questions 1 and 3), thematically, their content corresponded to the task examples with a negative sentiment (neg.sent.), suggesting that the proposed task classification is sentiment-independent.

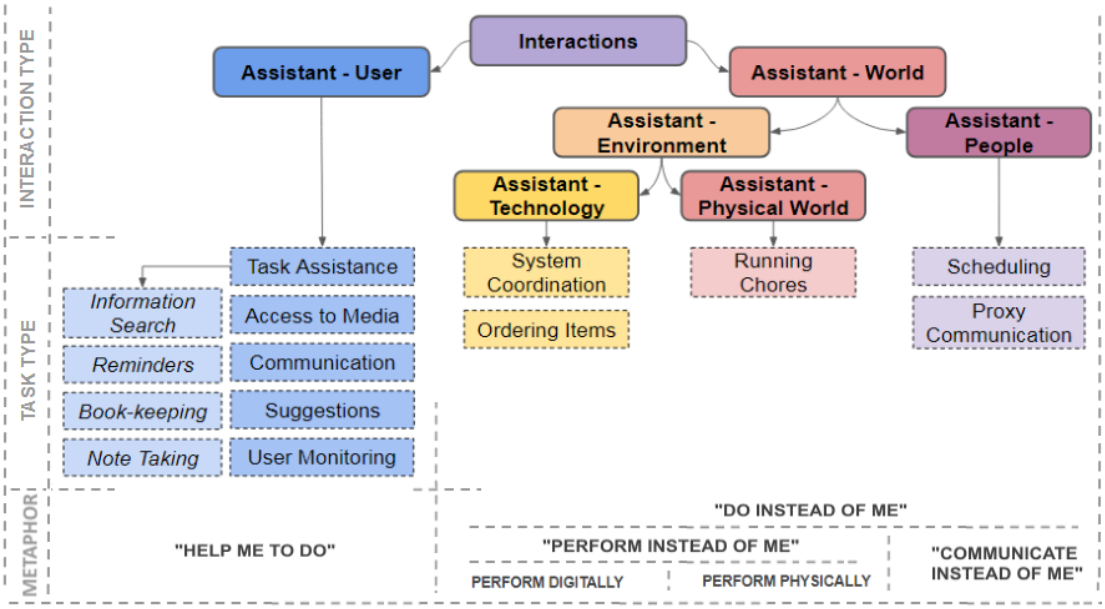

Figure 7.2: A classification of interaction types and corresponding typology of conversational agents' tasks.

Further analysis showed that these task categories form higher-level clusters that differ based on the combination of two implicit parameters: (a) the level of user's involvement in the primary activity and (b) the nature of interacting entities. The combination of parameters (a) and (b) configures the interaction process, hence, we refer to the resulting clusters as interaction types (fig.7.2, top). Our model includes six interaction types, hierarchically organized into a classification, and each type corresponds to an illustrative metaphor (fig.7.2, bottom). We first identified two "parent" interaction types - "assistant-user" and "assistant-world", according to the level of the user's involvement in the primary activity. To illustrate, consider the difference between task examples in the following pairs: *"tell me recipes" [P104]* vs. *"cook for me" [P99]*; *"add things to my calendar" [P64]* vs. *"make appointments for me" [P42]*; *"navigation suggestions" [P107]* vs. *"do lots of driving" [P53]*. The "assistant-user" and "assistant-world" tasks differ based on who is performing the primary activity. In other words, provided examples are either "help me to do", or "do instead of me" tasks. The "assistant-user" interaction corresponds to the "help me to do" metaphor (fig.7.2, bottom), and encompasses tasks intended to facilitate the user's main ac-

tivity. The "assistant-world" interaction corresponds to the "do instead of me" metaphor, and includes CA tasks of performing the primary activity (i.e., the user is not directly involved in the performance). The primary activity within the "assistant-world" interaction happens between CA and the environment, or between CA and other people. Hence, we separate "assistant-environment" ("perform instead of me") and "assistant-people" ("communicate instead of me") interactions. Finally, the environment where the main activity is taking place might be either technological ("perform instead of me digitally"), or physical ("perform instead of me physically").

In this section, we show the relationship between interaction types and task activities by organizing the ten task categories according to the six interaction types.

7.4.1 Assistant-User Interaction

Tasks of the "assistant-user" interaction type take place between the CA and the user. This means that the CA facilitates the performance of the primary activity, while the user actually performs it ("help me to do" tasks). The assistant's facilitation differs based on the required level of "intelligent" processing (*"Play music"[P36]* vs. *"therapy from an assistant" [P78]*), and the temporal aspect of a task (*"remind me of appointments" [P194]* vs. *"suggest better lifestyle choices" [P82]*). Based on the combination of the content, involved device infrastructure, required level of intelligence, and temporal aspect, we identified five task categories within the "assistant-user" interaction type.

Task Category 1: Task assistance

This category includes tasks of providing general assistance in a primary activity performed by a user. Initially, we identified it as a stand-alone task category based on the set of generalized descriptions of task assistance (19 task ex. from 19 resp., all positive): *"help me out" [P27]; "assistance with problems" [P72]; "helps with everyday life. facilitate any task without error or judgment" [P24].* Further semantic analysis revealed the sub-categories within the "task assistance" group (fig.7.2) based on the temporal and intelligence aspects.

Information search: includes tasks of searching information to facilitate the user's performance of the primary activity (113 task ex. from 84 resp., 4 with neg. sent.). This, first, includes the direct information search online, both in general (e.g. *"search internet for answers" [P35]; "search anything up" [P9]*), and searching for specific information, most commonly related to weather and news (e.g. *"give sports scores" [P103]; "the weather, updates" [P29]; "news, trivia questions" [P98]*). This also includes more complex forms of

information search that requires secondary processing of information from several resources: *"Google a question and then provide a summary and possible options. i.e. I want to surprise my husband with a Steelers tickets. Which games work with our schedule? And what would the total cost be?" [P39]; "Give medical advice. e.g. I provide symptoms and it walks me through a diagnosis" [P97]; "articulate the main idea of a short article" [P161].* Some specified that the information should be provided in a form of directions on how to perform an activity: *"give me instructions to an exercise routine or cook" [P47]; "provide directions on how to do things" [P249]; "walk me through a tutorial" [P227].*

Reminders: includes tasks of facilitating user's activity through issuing reminders of upcoming events or required actions, as well as setting timers during the activity: *"Remind me at a specified time to do things, go to meetings" [P116]; "Set timers" [P23]; "tell me when to take something out of the oven" [P242]; "remind me when bills are due" [P264].* (39 task ex. from 34 resp., 2 with neg. sent.).

Bookkeeping: includes tasks of keeping records of the user's plans, transactions, etc., and managing and composing lists to facilitate corresponding required actions. Exemplar bookkeeping tasks: *"Compose a grocery list" [P77]; "To do lists" [P102]; "you should be able to tell it to create a list which you can then add to over time (maybe a week)" [P141]; "assisting with shopping lists and price comparisons" [P171]; "Making and maintaining lists (grocery). Storing recipes" [P175].* (26 task ex. from 19 resp., all positive).

Note Taking: this last sub-category includes tasks related to producing "in-the-moment" records of fuzzy information, and storing it for the later use. While in our data this type of tasks was mentioned relatively rarely (7 task ex. from 7 resp., 1 with neg. sent.), this is a separate sub-category due to the particular requirements for the input, processing, and information representation formats involved in the execution of these tasks. For example: *"Record verbal notes while I am driving" [P93]; "Taking personal notes" [188]; "transcribe/type oral communication" [P237].*

Task Category 2: Access to media

This category includes "in-the-moment" tasks (temporal aspect) that require access to media resources and output channels (infrastructure aspect). The main activity in this case is the consumption of the media, which CA facilitates by providing easier access to it: *"playing music/podcasts" [P244]; "Turn on radio and tell it which station to tune to " [P141]; "Reading to my kids" [P192].* Another aspect of CA facilitation is through translating the media into other formats: *"to read a book, article, or web page that is not already in audio format"[P218].* (41 task ex. from 32 resp., all positive).

Task Category 3: Communication

Tasks from the communication category address user's social needs, as well as facilitate a decision making process through discussions. This includes general conversations, conversations with a specific purpose, light social interactions, and playing games with a user. Consider the following examples: *"playing word/mind games" [P244]; "Have discussions, aid me in making decisions" [P8]; "If needed, carry on conversation" [P73]; "play games, tell stories, tell me jokes" [P249]; "Therapy from a voice assistant could help many people." [P78].* (17 task ex. from 13 resp., 1 with neg. sent.).

Task Category 4: Suggestions

This category includes tasks facilitating the main activity through processing information and providing advice. As opposed to the complex "information search" tasks, which provide "in-the-moment" assistance, the "suggestions" have a less immediate nature: *"teach me how to budget" [P163]; "adding free apps or services for certain parts of my lifestyle" [P85]; "Predict my next move so I don't make a mistake" [P173]; "suggest new products, better lifestyle choices" [P82].* (32 task ex. from 29 resp., 8 with neg. sent.).

Task Category 5: User Monitoring

This category refers to monitoring specific aspects of the user's life and producing corresponding reports or suggestions (40 task ex. from 28 resp., 17 with neg. sent.). The "monitoring", first, includes tracking user's physiological data: *"monitor physical activity (exercise, blood pressure)" [P82]; "monitor my glucose, blood pressure, falls, changes in heart rhythm" [P234]; "communicating anything it is able to monitor, i.e heart rate" [P223].* These tasks also include monitoring users' activities (mainly financial) or performance, and related outcomes: *"spending our budget alerts" [P82]; "Give me a total increase/decrease in my investments" [P116]; "Keep me on track" [P190].* Finally, it refers to monitoring the environment to facilitate users' activities, e.g. *"keeping track of food in the house, when it needs to be replaced; expiration dates" [P219]"*. Some also mentioned monitoring for safety issues, e.g. *"record when necessary (rape, murder, break in, etc. For possible voice identification)" [P143]; "if the assistants feel there is domestic violence, or a robbery." [P258].*

7.4.2 Assistant-World Interaction

While the "assistant-user" tasks are facilitating the performance of the primary activity, the "assistant-world" tasks contain the primary activity itself. In other words, instead of assisting a user's performance, an agent performs this activity on behalf of the user ("do instead of me"). The main activity in these tasks takes place between CA and the world, resulting in "changes" in the world, as opposed to the user's task-performance space. For example, compare *"make appointments for me (hair/nails/car service) using my contacts list" [P116]* vs. (*"find info like address and phone numbers" [P237]*). Based on the nature of interacting entities, we identified two sub-types of the "assistant-world" interaction: "assistant-environment" – manipulations with the environment ("perform instead of me"), including technological and physical environments; and "assistant-people" – tasks of social interactions on behalf of the user ("communicate instead of me"). We continue by discussing five task categories organized according to these interaction sub-types (fig.7.2).

Assistant-Technology Interaction

The "assistant-technology" tasks involve a CA performing the main activity on behalf of the user in the technological environment. The technological environment refers both to the device ecosystem and to the online environment, e.g., websites and mobile applications. This distinction is reflected in two task categories associated with the "assistant-technology" interaction: "systems coordination" and "ordering items".

Task Category 6: Systems Coordination refers to CA managing other systems on the user's behalf, mainly smart-home systems (light, appliances, etc.), or car systems, including actual driving: *"run functionality of a home" [P92]; "Home security, energy savings actions. To avoid thefts and energy waste" [P135]; "drive my car" [163].* While the infrastructure resembles "access to media", functionally, these tasks are performed not for, but instead of the user: *"starting coffee, running dishwasher, to make my life easier." [P248]; "make coffee, lock doors, things that would make life easier for me." [P129].* (155 task ex. from 79 resp., 20 with neg. sent.).

Task Category 7: Ordering Items includes organizing the process of getting merchandise items to the user. For example: *"ordering the dog food and sending it to me" [P133]; "Order groceries to be delivered, buy concert tickets" [P174]; "to order books on-line" [P99]* (48 task ex. from 40 resp., 18 with neg. sent.). An assistant performs the main activity (getting items), as opposed to facilitating the user in performing this activity (e.g. providing directions to the store), thus, these tasks represent the "assistant-world" interaction.

Further, these tasks are the "assistant-technology" interaction, since the main activity takes place through online resources and does not require direct interaction with people.

Assistant-Physical World Interaction

The "assistant-physical world" interaction corresponds to the CA performing the main activity on behalf of the user in the physical environment. From the infrastructure perspective, these tasks are, arguably, the most challenging to implement. In fact, this type of interaction exists separately from the "assistant-technology" interactions because current ubiquitous systems have not yet expanded into the corresponding areas of life. In our data, we identified only one task category within this type of interaction - "running chores".

Task Category 8: Running chores refers to daily routine tasks, such as cleaning, feeding pets, cooking, etc. In these tasks, the main activity is performed by an assistant and the result is "returned" to the user by the environment (a cooked meal, a clean house, etc.). Further, these tasks are the "assistant-physical world" interaction, because the main activity is performed through the manipulation in the physical environment, as opposed to digital, and does not require assistant's direct interactions with people: *"giving my cat his meds every day"[P219]; "clean my house, fold my laundry, do my grocery shopping."[P180]; "Picking my clothes, cooking" [P58]; "Household chores, cleans rugs, wood floors, wash dishes, dust, make dinner" [P202]; "taking out garbage, emptying cat litter" [P215].* (36 task ex. from 23 resp., 1 with neg. sent).

Assistant-People Interaction

Finally, the "assistant-people" interaction corresponds to CA tasks performed through direct communication with others, e.g., calling to make an appointment, or setting up a meeting with another person. Similarly to the other "assistant-world" tasks, the main activity is performed by the CA and the result is "returned" to the user by the world - e.g., a scheduled appointment. However, this type of interaction also differs from the other "assistant-world" interaction types. First, these tasks require a system to have human-like abilities for communication (as opposed to between-systems communication). Second, these tasks require particular attention to ethical and privacy considerations within the communication processes. Finally, from the user perspective, communication with other people through an agent invokes specific aspects of trust to this agent, touching upon one's identity management considerations. The "assistant-people" interaction encompasses two categories of task activities - "scheduling" and "proxy communication".

Task Category 9: Scheduling refers to arranging user's meetings, e.g. appointments, work-related and social meetings, restaurant reservations, etc. Thus, these tasks require access to one's calendar and contacting means (infrastructure aspect). For example: *"Making doctor's appointments; scheduling transportation for doctor visits" [P219]; "Make travel arrangements, coordinate and schedule meetings" [P174]; "make phone calls for me (schedule hair appt, oil change appt, make dinner reservations, etc)" [P272].* Note the difference between "scheduling" and "bookkeeping" tasks: while the first describes the main activity of arranging an event, the latter describes the facilitation of the main activity through adding to the calendar a prearranged event: *"Making reservations for me" [P193]* vs. *"Mark dates in my calendar" [P77].* (26 task ex. from 21 resp., 2 with neg. sent.).

Task Category 10: Proxy Communication refers to communication on behalf of the user, e.g. *"make calls to companies to do general business for me" [P96]; "taking some calls for me, leaving a message from the calls" [P206]; "Call places and ask questions for me (ex. call an ice cream shop and ask their hours)" [P42].* (67 task ex. from 53 resp., 12 with neg. sent.). While the "communication" tasks of the "assistant-user" interaction (sec.7.4.1) are mainly audio-based complex activities, the "proxy communication" varies in medium (audio/text/visual) and complexity (one action vs. full conversation), e.g. *"talking to the dog" [P244]; "Like something on social media" [P36].* This also includes contacting authorities in emergency situations: *"Call 911 when needed" [P230]; "emergency services. A disabled person lives in my home, and she cannot always make a phone call" [P199].*

7.4.3 Section Summary and Discussion

An inductive analysis of task examples revealed high-level functional differences between clusters of CA tasks. We consider them "functional", because these differences reflect the task performance processes through the combination of user's involvement in the primary activity and the nature of interacting entities. They are "high-level" differences, because they occur at a higher level of generalization than the specific task activities. We refer to these clusters as interaction types, since they describe configurations of the interaction process. Our model includes six interaction types, hierarchically organized into a classification system (fig.7.2). These types represent different configurations of the interaction process based on (a) the user's involvement in the primary activity and (b) the nature of interacting entities.

While, to our knowledge, no similar categorization was offered previously, such systematic categorization of interaction types contributes to the HCI community in at least three ways. First, it allows to define system requirements for each type of interaction and

corresponding tasks. This includes, but is not limited to, security and privacy considerations, requirements for the device infrastructure, aspects of the interface design, and others. For instance, tasks within the "assistant-people" interaction require considerations for managing an appropriate user information disclosure, distinct interruption/error recovery mechanisms, specifics of the voice interface, etc. However, these requirements would noticeably differ from corresponding considerations for "assistant-environment" tasks. Second, the separation of different interaction types based on the level of user's involvement and the nature of interacting entities provides an insight into the structure of the mental model of a task performance. The process of reconstructing this mental model suggests directions for both supporting further research, and informing the design of an appropriate feedback system [289]. Additionally, it informs a set of possible directions for the development of IPA technology in general, and the device ecosystem specifically. For example, the differences between technological and physical environments inform development directions for the ubiquitous between-system communication.

If interaction types reflect the high-level functional differences, second component of the model – the typology of tasks – deepens the detailed differences, and presents ten categories and four subcategories of tasks organized according to the interaction types they represent.

The task categories were formed based on the interaction of four parameters: content, involved device infrastructure, level of required intelligence, and the temporal aspect of the task performance. We inferred these four parameters from the results of the open coding data analysis, but they also find collateral support in previous research. For example, Catrambone et al. [83] discuss a number of task-related variables that might differentiate agents' tasks, including objectiveness (opinion-based recommendations vs. acquiring facts), which loosely corresponds to the level of required intelligence we used in our classification. Further, the framework suggests variables such as degree of time pressure and duration, which are comparable to our temporal factor of task performance. Similarly, the intelligence factor finds reflection in the way users test IPAs [416]. The typology of tasks was formed based on the similarities of collected data instances and cannot be necessarily collectively exhaustive or mutually exclusive. Thus, it should be viewed as a conceptual clustering rather than a strict classification. In fact, this highlights one of the contributions of the model. Although currently the identified task categories largely validate and expand on the set of previously reported tasks [416, 285, 70], the future development of CA technology presents an unbounded set of possible task activities. Research attention to the evolution of levels of fitness of newly appearing tasks to the suggested categories opens opportunities for monitoring and analyzing the dynamics of CA adoption patterns over time.

Since the classification of *interaction types* and the typology of *tasks* are categorizing elements of a different nature, we consider them as two phenomenologically separate clas-

sifications. However, the model presents them in close association: the task categories extend the interaction types. Through this representation, we highlight that both academically and pragmatically it is more advantageous to consider these two classifications jointly. First, top-down, it "populates" the interaction types with detailed task categories, illustrating the implementation directions for interaction types. Second, bottom-up, it contributes to the understanding of the mental model of CA tasks by connecting specific task activities to higher-level functional heterogeneity, and allowing designers to navigate between the different system requirements for different tasks.

These two classifications mainly reflect the static aspects of the work process. To capture the dynamics of task performance, we developed the third model component, which we present in the following section.

7.5 Task Execution Paths

As we analysed responses with negative sentiment, we found that a significant number of participants' concerns revolved around the notion of control over the resources dynamics in the task performance. In combination with the evidence of a weak mental model of the IPA systems [289, 68, 272], this motivated us to explore the dynamic aspects of the task processes. Following our findings on the CA interaction types, we anticipated potential functional differences in task execution flows for the "assistant-user" and "assistant-world" tasks (fig.7.2). We looked specifically at these two types of interactions, since they are "parent" interaction types in our classification and, thus, diverge both on the level of users' involvement, and the nature of entities involved. Based on these considerations, we constructed a schematic representation of two abstracted task execution paths with a focus on task progress and resource flows (fig.7.3). Paths 1 and path 2 represent "assistant-user" ("help me to do") and "assistant-world" ("do instead of me") interactions respectively, and differ based on which entity (agent or world) returns the result of the activity to the user.

Both task execution paths start by a user granting access or specifically passing resources to a CA. These might be financial resources (credit card information), personal information (social number, passwords, etc.), social information (contact lists, social media), private information (personal conversations, preferences, search history), access to the house and car systems, etc. Upon receiving it, CA passes these resources (or a subset of resources) to the world to execute the task, and, in correspondence to the passed resources, the world produces some result. After that, path1 and path2 branch out based on the further information flow. In the remainder of this section we describe and provide examples for both task paths.

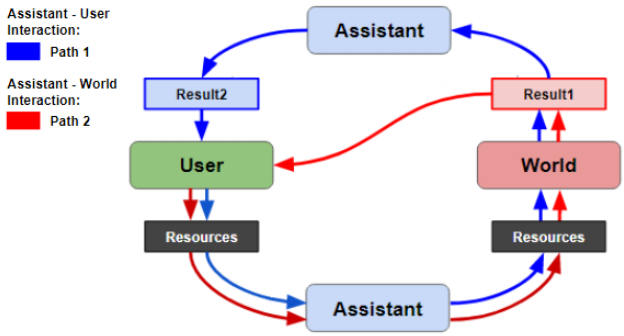

Figure 7.3: Abstracted task execution paths. 1: Assistant-User interaction type ("help me to do" tasks). 2: Assistant-World interaction type ("do instead of me" tasks).

7.5.1 Path 1: User-Assistant-User

This path abstracts the task flow when the result is returned by the assistant to facilitate user's performance of the main activity. After the user passes resources to the CA, the CA passes it to the world, and the world produces the result (result1). Then, result1 is returned to the CA, the CA receives it, and produces a new result (result2). The "assistant-user" tasks might or might not require secondary, "intelligent" information processing, thus, result2 might be either similar to result1, or refined through the "intelligent" processing. Finally, the CA returns result2 to the user. As an illustration, consider the following task description: *"Listen to our complaint/issue and offer suggestion" [P224].* An execution of this task, first, requires a user to pass to an assistant access to the private information (e.g. issues, complaints). The CA then processes this information and forms an appropriate search query (a subset of the user's resources), which is then passed to the world (Internet). Based on these passed resources, the world produces an intermediate result (result1), in this case, in a form of diverse pieces of information, and returns it to the CA. The CA further processes result1 into an informative summary or a suggestion (result2), and passes it back to the user to facilitate the user's problem solving process.

7.5.2 Path 2: User-Assistant-World-User

This path abstracts the task flow when the main activity does not require user's direct involvement, and the result is returned to the user by the world itself. After the user passes resources to the CA, the CA manipulates a required subset of these resources to the "world" to perform the main activity, and the "world" produces results, which are returned

directly to the user. Consider the following task: *"make calls for me to companies to do general business for me" [P96].* An execution of this task requires a user to supply the CA with access to contact and personal information (e.g. business objectives, phone numbers, etc.) and access to a communication medium. Then CA needs to process this information, form an appropriate "plan of action", and perform accordingly, by contacting businesses and communicating with them (revealing subsets of user's information). Finally, based on the passed information and agent's actions, the world produces some result which is returned directly to the user (i.e. the business is done).

7.5.3 Section Summary and Discussion

While the first two components of the model show functional differences between tasks, they do not reflect the dynamic aspect of the task performance. Meanwhile, our analysis showed a significant role of the sense of control over the resource dynamics in participant's cautions and negative sentiment. Thus, we created a schematic representation of two abstracted *task execution paths* with a focus on the resource flows and involved entities. Previously, Reis et al [383] also suggested a dynamic model for IPA interactions. Specifically, they developed a model for contextual interactions with the elderly, which includes the following steps: (1) user starts interaction, (2) the system identifies the user and acquires context data and personal information, (3) based on these data it generates an activity suggestion, and then (4) proposes this activity to the user. However, due to the design of this four-step model, it does not capture the differences between tasks, interaction types, or settings. In contrast, our diagram of task flows is designed to represent these differences, and to give a more granular outline of the dynamics of resources within tasks.

To summarize, the task execution flows augment the first two components of the model (interaction types and task activities) with the dynamics of task performance. Furthermore, it allows us to bridge the differences between "assistant-user" and "assistant-world" interactions to the areas of user concerns and cautions, discussed in the following section.

7.6 Problem Spaces

The thematic analysis of the second body of our data (perceptions of assistants' behaviors, usage expectations, and related rationale) revealed that negative sentiment mainly revolves around control over task execution and the flow of resources. We identified five clusters of concerns (problem spaces), and mapped them onto the task execution flows (fig.7.4). These

problem spaces include: unauthorized access to the user's resources; fear of being hacked; fear of agents acting on their own, without authorization; unauthorized passage of resources to the world; and lack of confidence in the quality of the task results. Corresponding concerns were present in answers for all six open-ended questions of the survey (tab.7.2).

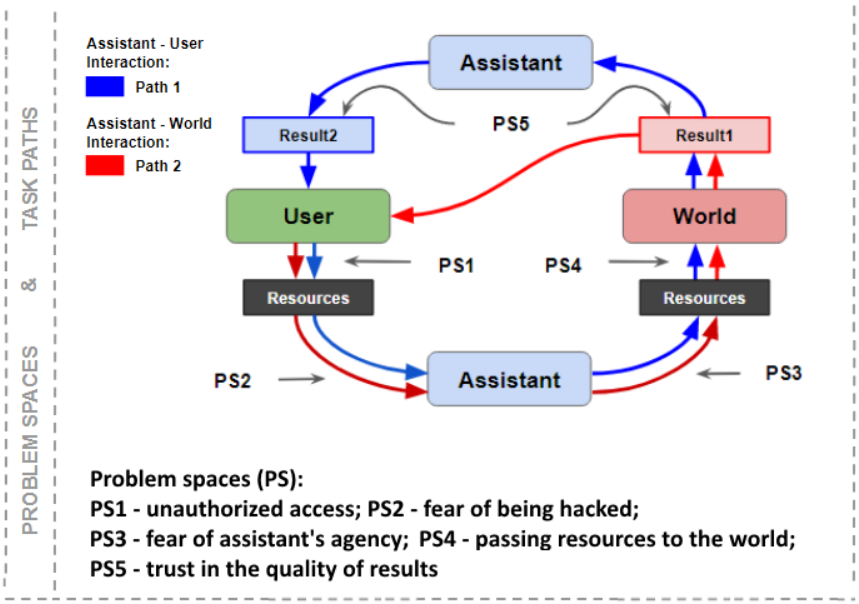

Figure 7.4: Problem Spaces (PS) mapped to the task execution flows, reflecting areas of users' concerns in relation to specific stages of task performance.

7.6.1 PS 1: Unauthorized Access to the User's Resources

The first problem space represents the beliefs regarding an agent's unauthorized and uncontrolled access to a user's resources, including financial resources, access to house systems, and information resources, such as personal, social, and private information (183 responses, tab.7.2). Consider the following examples of PS1 concerns: *"Eavesdropping, listening to conversations, private things" [P81]; "accessing information they were not told to" [P230]; "Access my camera without my permission" [P70].* This also includes unauthorized listening, which is tightly associated with participants' notion of lack of control of an overall working state of the system: *"If it was listening to me while I was not talking to it." [P64]; "If it turned on without me asking it anything." [P18]; "When they kick on without you asking them to." [P94].* Consequently, participants expressed concerns about potential unauthorized recordings: *"Turn on any cameras without being told, recording anything*

without being told" [P67]; "Save and record whenever with out my expressed permission." [P85]; "I do not need a voice assistant recording everything that is said inside my household" [P247]. Some authors ([68, 297]) report on "listening" and "recording" concerns separately, we, however, consider them as one cluster of concerns. Our rationale is that concerns about the information collection are, arguably, associated with the possibility of an agent doing something with it (intentionally or not), e.g. passing to third parties. PS1 refers specifically to the perceived ability of CA to get access to the resources, prior to any potential manipulations. Thus, we map PS1 to the beginning of the execution paths (fig.7.4), where a user is granting an agent access to the resources.

7.6.2 PS 2: Fear of Being Hacked

This problem space shows the lack of trust in the system's protection from malicious security attacks (72 responses). This leads to the fear that the information might be accessed by others through attacking an agent system, e.g. *"the possibility of being hacked" [P51]; "I would not trust the device with any personal information. I do not trust that anything I tell it or that it hears would not be hacked." [P141]; "Recording is something that could malfunction and if hacked would take away all sense of security" [P78].* We map this problem space to the part of the task execution paths reflecting the assistant's ability to "store" the user's information (fig.7.4). It is particularly important to consider this cluster of user concerns separately, since it reflects the lack of trust specifically in the system architecture as opposed to the trust to the system "intentions": *"constantly listening has to be included for assistants to work, but it also means somewhere in a repository my address, phone number, social security number, and everything else about me are being stored. That would not be a concern if cyber security was prioritized by companies, but, despite recent hacks, many companies have not bothered to ramp up security." [P261].* Hence, the system and interaction design should address user concerns from PS2 through the different mechanisms than, for example, when addressing fear of agents intentionally sharing the user's information externally (PS4).

7.6.3 PS 3: Cautions Towards Assistants' Agency

This problem space reflects the believed ability of an assistant to make independent decisions and act without specific request from a user (84 responses). For example: *"Have freedom to do anything without my approval/direction." [P46]; "Think on its own, make its own decisions." [P128]; "Make decisions for me about anything without asking for permission each time" [P85].* Concerns regarding the independence of agents' actions are tightly

Survey Question	Number of Responses				
	PS1	PS2	PS3	PS4	PS5
What would you feel comfortable with a voice assistant being able to do, and why?	8	1	1	3	1
What you would feel uncomfortable with a voice assistant being able to do, and why?	79	19	17	33	17
What behavior or traits of an assistant would negatively affect your trust?	18	20	25	23	48
What behavior or traits of an assistant would positively affect your trust?	64	23	34	56	41
Is there any particular rationale for how frequently you use voice assistants?	8	6	0	9	7
Additional comments on the topic of the survey.	6	3	7	9	0
Total:	183	72	84	133	114

Table 7.2: Articulated concerns in each problem space per survey question

associated with the sense of control over making a user's resources (especially financial) accessible to the agent: *"Decide to order products or services because I am out of the product" [P252]; "Sending messages from me without my specific [command]. I don't want my refrigerator send a grocery list to the supermarket. I want to control that function." [P176].* While PS1 also includes concerns about some unauthorized actions (unauthorized access to the resources), functionally, it reflects only the process of obtaining these resources. PS3, in contrast, refers to the fear of unauthorized use of these resources. Hence, we map PS3 to the part of the flow where an assistant is initiating a task, and manages resources required for the task performance (fig.7.4). The distinguishing feature of PS3 is that concerns are specifically related to the control over the initialization of actions, as opposed to the notion of agents having malicious or detrimental intentions (PS4).

7.6.4 PS 4: Unauthorized Passage of Resources

This space encompasses the apprehensions of CA intentionally sharing user's resources, especially information, with third parties without explicit authorization (133 responses). Consider the following examples: *"Telling someone my personal info" [P28]; "Sharing any data or information with anyone except direct family or given permission." [P67]; "Sharing anything I had not authorized with others, including other computers." [P39].* These concerns are often associated with the notion of an assistant "working" in the interest of

a third party, rather than in the user's interest, and thus, having "malicious" intentions: *"reporting back private conversations I've had in my home" [P25] ""spying" on me/us by reporting anything to anyone without my knowledge and explicit consent." [P119]; "Listening in and transmitting information to a company's servers." [P155].* Both PS2 and PS4 cover third parties getting access to the user's information. However, the important difference between them is in the element of the perceived agent's intention and role in the information leakage. PS2 refers to the lack of trust in the system security protection, and PS4 reflects believes that an agent purposefully passes the information to a third party ("not secure storage" vs. a "spy"). As an illustration, consider the difference between the following: *"Having too much private/personal information (SSN, ID, credit card information, passwords, emails), because a hacker or thief could steal identity, or kidnap a loved one." [P177]* vs. *""spying" on me/us by reporting anything to anyone" [P119].*

7.6.5 PS 5: Trust in the Quality of Results

The last problem space contains concerns with the agent's inability to accurately perform complex tasks (114 responses), e.g. managing home systems, communication, assessment of the situation, etc. These concerns vary based on the believed underlying reason for potentially unsatisfactory results. For instance, some concerns are related to the interruptions or malfunctioning of the system: *"Controlling standby electrical power equipment, because if system is interrupted it could result in frozen pipes, floods, or food spoil losses." [P135].* Other concerns are caused by the lack of confidence in the agent's intelligence: *"complicated actions that could be impossible to get 100% correct in every case. They require my analysis, judgment, intervention, decision making" [P93]; "Testing alarms such as fire. I feel I should do that or have the fire department do it." [P101].* One particularly specified aspect of intelligence is CA's ability to identify contextually complex situations: *"I think humans should still be able to call/alert emergencies. Ironically, I have epilepsy and it would be *amazing* if the assistants *could* alert someone when I have seizures." [P158].* Since PS5 reflects concerns regarding the quality of outcomes of CA's actions, we mapped it on the task flows as corresponding to the results element (fig.7.4).

7.6.6 Section Summary and Discussion

The last component of the model is the classification of user concerns as five *problem spaces* mapped onto the diagram of task execution paths. The problem spaces are segregated based on the nature of concerns, and their correspondence to the system architecture components (input, storage, event initiation, data migration, output), to support

a homogeneous, process-oriented nature of the classification. The demonstrated associations between the problem spaces and their "place" in the task execution flow is the first systematic representation of user concerns in a dynamic context of a task performance.

The resulting problem spaces support and expand on previously reported results. Interestingly, some areas of concerns are explored more than others. For instance, the literature describes [68] user beliefs that CAs are always listening (PS1 in our classification), that the information is fully recorded and transmitted to the cloud (corresponds to PS2 and PS4), and that an agent's error might have noticeable consequences (PS5). Yet, there is little discussion of the perceived ability of agents to initiate a task on their own (PS3 in the model). The closest comparison can be drawn between our classification of problem spaces and the work of Manikonda et al. [297]. The authors investigated user privacy concerns regarding using the IPAs, and identified seven privacy issues: the device getting hacked, collecting personal information, listening 24/7, recording private conversations, not respecting the user's privacy, data storage repository, and "creepy" nature of the device. In reciprocal validation, some of these issues correspond to the problem spaces in this project. To address these issues, however, one needs to understand them functionally, e.g. why recording of private conversations is feared, or what does it mean to "not respect privacy". Our work extends on these descriptive results by suggesting a process-oriented classification of user concerns, linked to the task flows and the system architecture components. This process-oriented nature brings both academic and practical contributions. For instance, the association of user concerns to their "place" in the task execution flow informs the nuanced design of targeted solutions for particular problem spaces. It also contributes to the development of a mental model of CAs. Thus, it allows creating corresponding feedback and visualization mechanisms to address adoption obstacles [289], and an extrapolation of effects of system visibility on other ubiquitous technology suggests that an appropriate feedback might effectively guide privacy self-management decisions [367].

7.7 General Discussion

In this project, we explore the design space of general-purpose voice assistants. We discuss how both academic research and consumer adoption reports demonstrate the steady decrease of daily and monthly active use of smart speakers due to users' low functional dependency on CAs and the discrepancies between users' expectations and the reality of the system's operation. Furthermore, the non-use seems to be often rationalized by the "lack of usefulness" of CAs, and those who choose to use CA systems tend to carry out only a small number of available tasks, potentially due to the low discoverability of agents' skills,

underdeveloped understanding of CA systems, and concerns about system performance.

We argue that the complexity of these problems can be addressed through the development of an understanding of the CA design space. This understanding includes both appropriate tasks and functions and factors that influence design acceptance such as users' beliefs and concerns. Correspondingly, this chapter reports on the initial exploration of this design space through user-centered scenario-based envisionment probes. Specifically, we present an empirically grounded four-component conceptual model of the CA work process which includes users' functional needs (tasks), the underlying interaction model (interaction types), as well as users' expectations and beliefs regarding CA technology (resource flow and problem spaces in task execution paths).

The first component – the hierarchically organized classification of six *interaction types* – represents different configurations of the interaction process based on (a) the user's involvement in the primary activity and (b) the nature of interacting entities. The conceptual difference between the two parent interaction types – assistant-user interaction ("help me to do") and assistant-world interaction ("do instead of me") – opens an interesting question for the fundamental understanding of the interaction design. Specifically, while we commonly refer to conversational agents as voice *assistants*, the breadth of "do instead of me" interactions suggests an unexplored tension between *assistant* and *collaborator* roles of an agent. At the same time, each of these roles, arguably, implies a number of distinct interaction specifics. For instance, these roles would differ in the expected user-agent balance in the decision-making processes [395, 149], might affect the mechanisms of the sense of ownership over the technology [248], would influence the anthropomorphization of agents [254], would inform an appropriate user-agent coupling in task performance [327, 359], and might require different trust dynamics toward an agent and their performance. Furthermore, from the perspective of design and development implications, each of the six interaction types has it's own architecture requirements. Additionally, the system of interaction types directly informs the structure of skill templates and suggests how to approach users' mental model of interaction with CAs.

The typology of tasks – the second component of the model – reflects the detailed differences between types of *task activities* with CAs envisioned by participants. Formed based on the interaction of content, device infrastructure, level of required intelligence, and the temporal parameter of task performance, the task categories contribute to our understanding of users' functional needs, and how to improve the discoverability of new skills through building suggestions for tasks based on the users' situational expectations. Interestingly, most of the commonly adopted skills are associated with the assistant-user ("help me to do") interaction type, e.g. playing music (access to media), information search, reminders, creating lists of actions (simple bookkeeping), and taking notes. In

contrast, while assistant-world ("do instead of me") tasks are often technologically possible or, even, already implemented, their current implementation does not seem to trigger common adoption. The significant weight of these tasks in the dataset of user-envisioned actions suggests that there is an interest and a functional need for such skills. The low rates of their adoption, then, might be a sign of a design misrepresentation of the types of tasks, such that they do not properly fit or address the users' model of CA interaction. This further highlights the contribution of our model which considers the task clusters in relation to the interaction types, uncovering the corresponding users' concept mapping. Another reason for lower adoption of "do instead of me" tasks might be that these skills require targeted design solutions that differ from "help me to do" tasks. For instance, they require distinctly different feedback support from the system to develop and maintain users' appropriate sense of control over the recourse flow in the task execution paths.

To reflect this, the third and forth components of the model – the *abstracted task execution paths* for "help me to do" and "do instead of me" interactions and five *areas of user concerns* mapped onto the task execution paths – reflect the dynamic aspect of task performance expectations. While the first two components of the model show functional differences between tasks, the dynamic aspects of the third and fourth components provide additional details on desired functionality and factors that influence adoption of CAs. Specifically, mapping of the problem spaces onto the task execution paths and related resource flows illustrates how the procedural nature of the task performance affects the differences in requirements for "help me to do" and "do instead of me" tasks, and allows us to associate user concerns regarding the potential pitfalls to the specific steps of task performance. Thus, while other research has looked at different aspects of these discussed problem spaces, the lack of their overall analysis in the context of dynamic task performance made it challenging to contextualize past work (e.g. user concerns) in the systems design.

We see the main academic significance of our model in the opportunity to initiate generative research. Given that CA systems are rather new for everyday use, it is expected that current research on their adoption is predominantly observational and information-gathering. However, building upon the rapidly growing body of descriptive research, it seems timely to begin directing effort towards establishing underlying mechanisms and developing a theoretical understanding of interactions with ubiquitous general-purpose voice-interface technology. The organization contributes both to supporting the visibility and cognitive model of the system, and supporting system architecture by allowing navigation within the types of interactions, associated tasks, and problem spaces. Finally, each model component and their joint composition are intentionally structured in a way that allows further expansion and augmentation. This permits both a much needed flexibility in the rapidly-evolving field of CA research, and an opportunity to analyse changes in the

field by considering the evolving fitness to the components of the model.

7.7.1 Limitations and Future Work

The study was designed to explore users' perceptions and beliefs regarding appropriate tasks and interactions with conversational agents in home settings. Participants were prompted with a hypothetical scenario of an adoption of a conversational agent system in their home with no technological limits to its functionality. There is a number of advantages of the thought experiments format, including the ability to unbinds participants' conception and reasoning, and to overcome the limitations of purely descriptive data on already existing practises.

However, one challenge with any online form of data collection is that there are well-known limitations. For instance, while some participants provide in-depth elaborated responses, others confine themselves to more condensed answers. Unlike interviews, the format of the data collection does not allow follow-up to further expand our understanding of participants' responses. Thus, online data collection provides high replicability, but can lack some of the contextual richness of the narrative. In our analysis, we address this issue by cross-validating our results with findings from previous research. Furthermore, the ability to capture information from a large number of participants (in our study 272) allowed us – possible because of the design of our survey – to capture a large set of elaborated responses from which to synthesize theory.

In this work, we presented the results of an exploration of CA design space, including users' functional needs, the underlying models of interaction, and users' expectations and beliefs regarding CA technology. Building on these results, future work should explore how we can scaffold use of CAs to address discoverability issues by leveraging the users' understanding of the system of tasks and interaction types. Furthermore, while the two abstracted task execution paths outline the high-level differences for specific interaction types, understanding of these paths in greater detail would advance the understanding of users' mental model of the CA work process and would strengthen feedback design. Finally, the last component of the model – the system of problem spaces mapped onto the execution paths – suggests how each problem cluster should be addressed differently in providing users' with the sense of control over their resources. The future research should further explore the specifics of these differences and the effectiveness of particular design solutions for each of the problem clusters.

7.8 Conclusion

This project contributes to the systematic understanding of the functional differences between CA tasks in association with the user's concerns about this technology. Specifically, we present an empirically grounded conceptual model of the CA work process, which includes four components: (1) a classification of interaction types, (2) a corresponding typology of CA tasks, (3) two abstracted task execution flows associated with the interaction types, (4) a classification of users' concerns mapped onto the task execution flows. This model provides direction for the further structuring of CA work processes, and highlights the importance of the dynamic association between CA task processes and users' conceptualizations.

To this point in this thesis, we have explored communication through, around, and with technology through six individual projects. In the remainder of this thesis, we will explore how we can synthesize our learning to better understand the ephemeral nature of communication that is mediated by or directed toward technology.

Part IV

Discussion

In this thesis, we consider a system to be any computational technology that participates in user-initiated information communication, whether the user communicates through the system (with others via a system-supported communication channel), around the system (with others, leveraging information provided by the system), or with the system (information is destined for the computational technology itself). We use the generic term "system" to articulate the varying way that computation participates in the information sharing, i.e. as direct channel, side channel, or recipient.

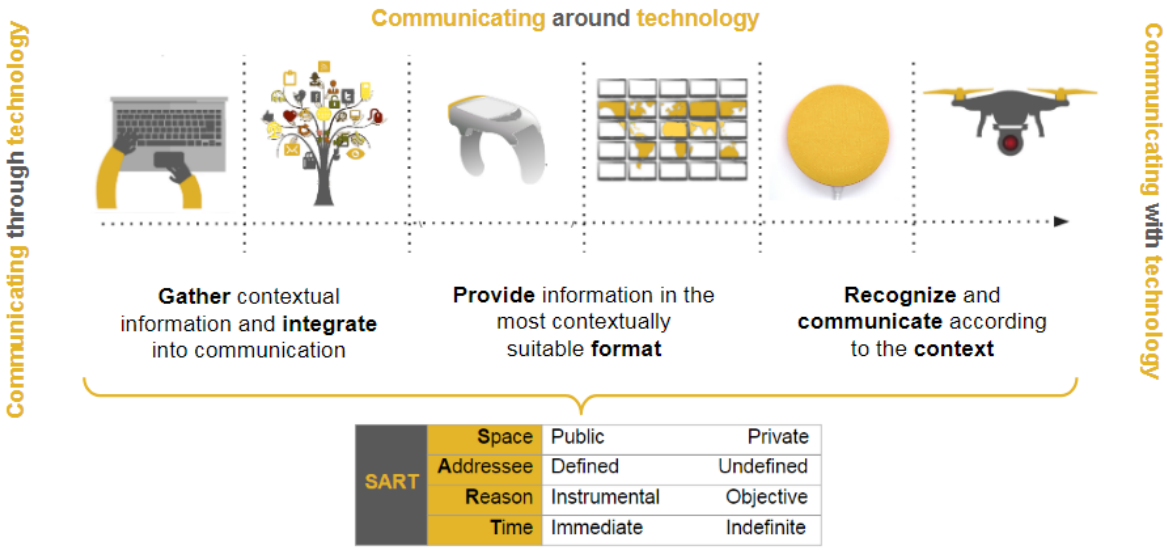

Figure 7.5: Summarized findings presented in this thesis.

We advocate for the need to manage information shared by a user according to the user's information sharing intent, which in turn is defined by the information sharing context.

We suggested that, depending on the system's position between a user and an intended receiving agent – whether communication happens through, around, or directly with the system – the system would have different patterns of operational adaptation to communication context (Fig.7.5). Specifically, we demonstrated that when communication occurs **through the system**, the system's general operation adaptation pattern is *to gather communication context unavailable to the communication endpoint* and *to integrate it into information communication* (Part I); when it happens **around the system**, the system should adapt its operations *to provide information in the most contextually suitable format* (Part II); and when a user communicates **with the system**, the role of the system is to recognize and *"match" this context* in communication with the user (Part III).

Furthermore, the implementation of these operation patterns to achieve intent-grounded information handling requires a system's ability to infer the user's sharing intent. This ability is predicated upon our understanding of how the user's sharing intent is affected by the information sharing context. In Part IV of this thesis, we first discuss the role and conceptualization of context in light of existing literature (chapter 8, sections 8.1 and 8.2). Then, through a cumulative analysis of findings from individual projects, we identify a set of **contextual factors** (chapter 8, section 8.3) affecting user's sharing intent **across the system types** (Fig.7.5). Finally, we conclude by outlining directions for future work and summarizing work in this thesis (chapter 9).

Chapter 8

Context and SART Framework

In this chapter, we first discuss the role of context in the user's information sharing intent in light of literature in different domains of computer science and illustrate the lack of actionable links between this research and design of context-aware communication systems (section 8.1). To address this issue, we proceed by discussing how context can be understood by the system and what systems can do in response to it (section 8.2. Thought this discussion, we positioning our understanding of contextual factors in the frame of this thesis, and, consequently, we present four contextual factors, synthesized in an early design framework (SART) based on our work and evidence from literature (section 8.3).

8.1 Role of Context in Information Sharing Intent

The effects of context on the user's communication and information sharing have been of interest to several domains within computer science, e.g. the development of assistive technology [438, 293, 182, 442], wearable sensors and tracking technology [509, 403, 475], smart homes [31, 531, 449], or computer-supported collaboration [407, 418, 488, 26, 359, 476], to name only a few. Research on computer-mediated communication has extensively discussed the role of users' shared situational awareness – understanding of the communication context of each other – in establishing effective communication [199, 190, 48, 200]. Within this domain, research has mainly focused on understanding the limitations of the communication channel on providing context awareness to the communication endpoints [57, 56]. Similarly, my own research, early in my academic career, explored impression formation and intent interpretation through different communication channels [256, 255].

Perhaps the most extensive attention to the role of context in information sharing has been explored within the privacy domain, which, first, suggests that the perception of information as private is context-dependant [436, 296, 355, 356], and second, demonstrates that information sharing is a contextual decision rather than a fully rational act [506, 456, 525]. Specifically, while the *privacy calculus theory* suggests that individuals calculate between the expected loss of privacy and the potential gain of disclosure [237, 506], the perspective of a benefit seems to be heavily biased [5]. Sundar et al. [456] showed an important role of affective impressions in priming sharing decisions and, based on these results, suggested that, while sharing attitudes might be produced by systematic, logical processing (context-independent), actual sharing behaviors tend to be regulated by heuristic processing at the moment of sharing (context-dependent). These differences in underlying process lead to the discrepancy between privacy concerns and actual sharing behavior [302, 386], labeled the "privacy paradox" [43, 33, 194, 458]. Furthermore, the effects of diverse social biases on information sharing decisions challenge the consideration of the user's agency in information sharing processes. Similarly, from sociology, *structuration theory* [176] brings together human agency and the social structure that affects and is affected by this agency [237]. Adopted for digital privacy [525], structuration theory views privacy decisions to be heavily influenced by contextual factors, as opposed to an objectively free user's agency in the sharing decision.

In fact, multiple privacy theories exist that acknowledge or consider the effects of context on information sharing decisions [436, 296, 355, 334]. In essence, these theories suggest that a specific piece of information might be considered private in some contexts, but public in other contexts. Thus, content is secret not on its own, but in a particular information sharing context which impacts the types of privacy rules that individuals use in managing information [355, 356]. For instance, Solove [436] suggests that the context of a given privacy problem should guide our understanding of privacy in a particular situation, which determines whether or not information is private. Mai [296] proposes that instead of considering informational privacy from perspective of controlling, limiting, and restricting access to of information, it should focus on a regulation of the use, analysis, and interpretation of personal information, since this information itself is contextual and as such open for interpretation. Nissenbaum's theory of contextual integrity [334] states that users control their privacy needs depending on the respective context-specific substantive norms that constrain what information websites can collect, with whom they can share it, and under what conditions it can be shared.

Suggesting potential contextual influences on the information sharing intent, one of the most famous privacy theories, suggested by Westin [499], considers privacy both as a non-monotonic function (people can have different degrees of privacy) and as a dynamic process

(it is changing to serve momentary needs and role requirements). Similarly, Altman's privacy theory suggests that privacy regulation is a process of regulation of interaction with others, which changes in response to one's internal states and external conditions [13]), becoming "the selective control of access to the self" [12]. Importantly for the purpose of our discussion, Altman views privacy as inherently a social process, which is determined by the interplay between social (communication) actors, their environment, and related temporal aspects [13]. Finally, Petronio's communication privacy management theory [355], constructed specifically for interpersonal computer-mediated communication, extends Altman's theory and considers privacy from the perspective of personal and collective "boundaries". This theory is particularly valuable in application to the system's inference and support of users' information sharing intent, since it suggests that information disclosure is regulated by rules. This includes rules of boundary permeability (how much others can know about the information), linkages (who else can know the information), and ownership (how much control co-owners within collective boundaries have over the information), and rules might be implicitly expected based on a person's assumption that others (including the system) share their understanding of these rules [94, 355]. For a comparative review of Altman's, Westin's and Petronio's theories of privacy see [298].

Taken together, this research suggests that information sharing decisions are indeed guided heuristically according to sharing context. However, while privacy research discussed above forms an essential basis for understanding of the role of context in information sharing, it provides little functional insight in how context can be used to infer users' intent. It shows that decisions are affected by affective priming [456], social interactions [13, 355, 334] and cultural norms [525, 13, 355, 334], situational needs [499], and environmental and temporal aspects [13], but lacking actionable evidence on identifiable values for these influencing aspects that could be used for sharing intent inference across the systems.

In the remainder of this chapter, we attempt to approach this gap by synthesizing findings from our work and from literature in an early design framework of four contextual factors affecting information sharing intent (SART framework, chapter 8). Since the system's context-based inference of sharing intent requires an understanding of how this context can be approached by the system, we first define the scope of our context conceptualization. Reflecting the complexity of the phenomena, we positioning our understanding of it according to the dimensions of the source of contextual factors, the immediacy data, and functionality. We then discuss each of the four contextual factors in SART, grounding it in both our individual research projects and in related findings and theory from literature.

8.2 Understanding Context

In the field of HCI, we refer to systems which are able to adapt their operations to the relevant context without explicit user intervention as context-aware systems [29, 464]. Correspondingly, context-aware communication systems are expected to take into account the situation of the parties and adapt particular steps of communication according to users' contextual needs [469]. The importance of context-awareness has been recognized by communication systems researchers [188, 412]; however, the development of a systematic approach to context aware communication systems is predicated upon understanding of (1) what is it that the system adapts to (*what is context for the system?*), (2) how does the system adapt to it (*how does context affect the system's operation?*), (3) and why does the system adapt to it (*what is the function of these systems?*).

What is context for the system?

In context-aware systems design, the conceptualization of context was developed alongside technological advances. Early definitions of context [410, 413, 120, 239, 1, 66] were mainly component-based, focusing on listing specific types of information necessary for different applications [464], which predominantly revolved around using sensible environmental data [489, 411, 2, 150, 454]. Soon, however, it became apparent that a component-based approach to context, leading to a separation of the system's actions from the underlying context that causes those actions [41], is not reflective of the complexity of the phenomenon.

This led to a growing appreciation of the heterogeneous nature of contextual data and more complex consideration of what constitutes contextual factors for the system.

For instance, researchers started reflecting the heterogeneity of context through the dichotomy of "physical" (external) versus "logical" (internal) context [192, 361, 29, 208, 469, 413, 41]. In line with past work, the internal context describes user state, experience and activity [192, 29], and external context refers to environmental state, i.e., location, light, movement, etc. [29, 361, 192].

Within this dimension, contextual factors in our work represent **"logical"** context, since they are internal to the user, e.g. their believes and expectations, rather than the user's physical environment. Note, that, although the nature of the considered contextual factors is **"logical"**, the raw, low-level data used to infer them [192, 413, 41] can be both internal (e.g. network of user's accounts) and external (e.g. time or physical proximity).

Furthermore, the evolving understanding of context facilitated research interest in context modeling [413, 507, 29] and using aggregated low-level sensable context data to build

higher-level context infrastructures [469, 447, 29, 413]. Here, the dichotomy of "captured" versus "inferred" context reflects the level of processing required for a system to consider contextual data as a contextual factor. For example, several communication systems have focused on exploiting the contextual factor of "busy-ness" (or "interruptability") of a user. Fogarty et al. [158] presented a system that modeled availability for communication using the built-in microphones of laptop computers to sense nearby speech, in combination with location, computer, and calendar information. Nagel et al. [321] explored the use of context information to signal the user's availability to close friends and family at home [322]. ConChat [374] used environmental sensors to allow users to infer during a conversation what the other person is doing and what is happening in his or her immediate surroundings to assess the expected speed of message exchange. Finally, Munõz et al. [317] developed a system for a hospital setting that used a set of circumstances (location, time, roles, and device state) to identify when to send messages to the recipient.

Contextual factors in our work represent **"inferred"** context, since they depend on, but extend beyond, data directly accessible for the system to capture. While, typically, any **"logical"** contextual factor would have to be inferred, the responsibility for such inference would often lay outside of the system's operation. For instance, the Community Bar system [462] shared screenshots of the user's display with their co-workers to provide an awareness of the user's work process, and the Montage system [459] allowed users to remotely briefly activate the webcam on another user's computer to assess whether they are busy.

At the same time, the domain's consideration of context began to evolve from the representational approach of a situation description, toward the interactional approach concerned with what makes information contextual [3, 41]. This interactional approach defines context by situated practice – how context is used to work on a particular task [217], and thus takes the system's design focus from the ability to sense the environment [410] (what the system uses as context) to its ability to functionally support the task based on contextual information [122, 361, 74] (what the system does with context). For instance, Dourish [130] describes these differences in conceptualization of context as a distinction between positivism – a component-based [41] environment sensing approach, looking at context as a representational problem [11], – and phenomenology – the recognition of context as an interaction problem [11], particular to each instance of activity and defined dynamically. Dourish [130] uses the concept of practice to link the user's action and its meaning and suggests we view context not as a set of descriptive settings, but as a dynamic form of engagement with those settings. Similarly, Coutaz et al. [108] distinguished between a more complex and flexible understanding of "context-as-process" and simpler understanding of "context-as-state".

The interactional approach to context as dynamic engagement with the situation [130,

361, 74] also allows us to approach our second question – how does the system adapt its operation to context.

How does context affect system's operation?

In essence, the conceptualization of context as a dynamic practice augments an "objective" view on users' actions by introducing the consideration of the meanings (interpretations and consequences) of users' actions for themselves and for others; users create and communicate these meanings though technologies they adopt [130]. The non-system-specific answer to the question of how context should affect the system's operations lies in adaptation of operations to accommodate the meaning of the user's actions. As an illustration, consider specific examples of adaptation of operations, e.g. providing recommendations selected according to the interpreted meaning of a user's search (context-aware recommendation systems) [197, 8, 472], adapting configuration of user interface according to the interpreted meaning of a user's actions (context-aware adaptive interfaces) [437, 333, 139], or adapting navigation to accommodate interpreted meaning of a user's movements [80, 400].

Context-aware communication systems can be viewed as adapting their operations to accommodate the interpreted meaning of users' communication actions. In our work, we enhance this notion by demonstrating that, depending on the system's position in the communication process – whether communication happens through, around, or directly with the system – the system would have different patterns of required operational adaptation to communication context. Specifically, when communication occurs through the system, the system's general operation adaptation pattern is to gather communication context unavailable to the communication endpoint and to integrate it into information communication (Part I); when it happens around the system, the system should adapt its operations to provide information in the most contextually suitable format (Part II); and when a user communicates directly with the system, the role of the system is to recognize and "match" this context in communication with the user (Part III).

What is the function of context-aware systems?

Finally, since context-aware systems adapt their operations to accommodate meanings of diverse user actions, the last big question is "why", i.e. what is the purpose of adapting the system's operations to accommodate the meaning of users' actions.

While each particular context-aware system has some targeted functionality (e.g. providing recommendations [197, 8, 472], adapting the interface [437, 333, 139], or providing

information on the user's availability [52, 158, 532]), the overall purpose of their adaptations tended to revolve around providing "convenience" of use [29]. We refer to this focus of functionality as "context for convenience" to distinguish it from a focus on supporting context-dependent meaning of an object or event – "context for interpretation". The use of context to adapt the system's operations to support interpretation has been actively exploited in other areas of computer science (e.g. vision systems for recognition [468, 339, 275], natural language processing and generation [171, 75], or activity recognition [453, 338]). However, within context-aware systems design, the functionality focus on *"context for interpretation"* has received very limited attention. The functionality of the contextual factors suggested in this work seeks to support interpretation processes in their own right; thus, functionally, we refer to the contextual factors of interest as the *"context of interpretation"*.

The overall purpose of context-aware systems that support context for interpretation is to support transferable context-dependent meaning of an object, event, or any other data. Given that the need for the system's ability to recognize users' sharing intent is motivated by the need for intent-grounded information handling, we argue that the system's functionality should be considered as focused on "context for interpretation".

The use of contextual factors to adapt the system's operations to support *interpretation* has been actively exploited in different areas of computer science (e.g. [468, 339, 275, 171, 75, 453, 338]), but not typically in the design of context-aware systems. The few examples of systems with functionality focused on supporting information interpretation come from recent developments in accessibility. For instance, Stangl, Morris and Gurari [438] explored visually impaired people's experiences with descriptions of digital images from diverse contexts, including news websites, social networking platforms, eCommerce websites, employment websites, online dating platforms, productivity applications, and e-publications. They found that image description preferences vary based on the source where digital images are encountered and the surrounding context. Similarly, in an earlier study, MacLeod et al. [293] demonstrated how visually impaired people experience automatically generated captions of social media images and fill in details to resolve differences between an image's context and an incongruent caption. Grayson et al. [182] introduced a computer vision-based AI system that provides people with vision impairments dynamic, in-situ information about the location, identity and gaze-direction of people nearby. Authors discuss how this information supports contextual interpretations and enables visually impaired users to develop their communication skills and be more confident in their social interactions. Furthermore, Stearns and Thieme [442] proposed a computer vision system that can automatically detect people in dynamic real-world scenes, enabling people with vision impairments to have more awareness of the communication context in the environ-

ment.

To summarize, in this section we discussed what context is for the system, how context affects the system's operation, and why the system adapts to context. Through these three questions, we defined the scope of our understanding of contextual factors and positioned it according to three dimensions: the source of contextual factors ("logical" factors), the immediacy of contextual data ("inferred" factors), and the "context functionality" (context of "interpretation"). In the following section, we discuss what particular contextual factors seem to affect users' expectation on interpretation and use of shared information (users' information sharing intent).

8.3 SART

One challenge with designing computational systems to support communication (whether users communicate through, around, or with systems) is the disparity between users' expectations on the information interpretation and use (i.e. users' sharing intent) and the reality of information retrieval by the system. To begin to address this disparity, in this section I synthesize findings from the six projects presented and identify a set of four contextual factors that seem to affect users' information sharing intent.

These four contextual factors are presented as a framework, which we refer to as **SART** according to the included factors included **space, addressee, reason and time**, where each context factor is defined through a descriptive dichotomy (Fig. 8.1). In a further effort toward operationalization of these factors, we define each of them as a continuum through a dichotomy of descriptive "values", and, based on our initial analysis, outline observed mechanisms of their influence on information sharing intent.

SART	Space	Public	Private
	Addressee	Defined	Undefined
	Reason	Instrumental	Objective
	Time	Immediate	Indefinite

Figure 8.1: SART, early design framework

8.3.1 Space

The **space factor** is represented by the continuum from "public" to "private" space, and refers to the perceived breadth of the audience for the shared information.

In our research, we have observed the effects of space interpretation, for example, in users' practices and perception on online reputation content. Specifically, **project 1** (chapter 2) revealed that both searchers and searched individuals want to prioritize and categorize content to constrain the accessible information to the specific goal of their search – targeted narrowing of the breadth of the audience. Inability to do so leads to the user's strategic limitations on the information that they initially provide online; however, such limitations harm online presence in other aspects of the user's life online.

Another example of the effects of space interpretation was observed in relation to sharenting behaviours (**project 2**, chapter 3). In particular, within the intended communication context with a parent's social circle, the protagonist tends to be assigned to a parent. However, when the audience becomes less defined for a sharing parent – the present and future "unintended" audiences – the expected shared communication context is lost, and the protagonist is more likely to be assigned to a child. Correspondingly, a parent's considerations of their "intended" (narrow and "controlled") audience tends to provoke higher readiness to share, loosening sharenting practices, while perceptions of potential broader audience provoke stricter sharing considerations.

Similar effects of space interpretation on a user's willingness to share information can be observed in parents' underlying motivations for using child monitoring technologies (**project 4**, chapter 5). While such information about their child is seen as highly sensitive and motivates increased attention to security and privacy consideration, its intended use within the private (even rather intimate) relationship between parent and child increases parents' willingness to collect such information through technological mediation.

The opposite effect of space interpretation appears in privacy considerations in interactions with conversational agents (**project 6**, chapter 7). For instance, the interpretation of space as private is compromised when a user perceives the agent as intentionally "spying" to share a user's information with third parties. Furthermore, these effects become particularly relevant when an agent performs direct communication with others (e.g., calling to make an appointment), invoking specific expectations for an agent to make a decision on what information can be provided to other people, and thus, to differentiate between a user's "private" and "public" information.

Gal [168] suggests that the use of distinction between public and private should be analyzed as a communicative phenomenon to capture systematic logic and underlying so-

cial reasoning about it. In other words, perceptions of something as public or private is dynamic and relative, so that any notion of "public" or "private" makes sense only in a paired opposition [495]. For instance, one's home can be considered as a private space when contrasted with the neighborhood, but have public and private spaces within itself [260]. This relative nature of private/public assessment can also be seen, for example, in the perceptions of personal devices. Chin et al. [96] explored user attitudes toward privacy of smartphones in relation to attitudes toward more traditional computing systems, e.g. laptops. They demonstrated that users are generally less willing to perform sensitive tasks on their phones, particularly due to the expectation that they do not feel their information properly protected (trust in Wi-Fi or 3G networks), while desktop computers were perceived as more private. On the other hand, in relation to public displays, smartphones tend to be viewed as more private, which is actively exploited in the design of personalized information handling in interactions with public screens (e.g. [79, 30]).

However, interpretations of space as private or public, occurring in the physical world, do not seem to naturally transfer to digital spaces. For example, Fuberek et al. [189] investigated how undocumented immigrants use technology, their risk perceptions, and protective strategies around their vulnerability, and found that, while this population would act to address offline threats, the behavior did not translate to their online activities. Similarly, Witmer [508] explored users' engagement in risky forms of online interactions and found that nearly half of participants believed their messages to be private and perceived risks of such interactions to be low. Thus, while the physical world, time, and space allow people to separate different social contexts, online communication does not seem to appropriately support privacy barriers [128] and selective sharing mechanisms [222], although overall evidence tends to suggest that the user's willingness to share personal information decreases with the number of recipients [414].

Burkell et al. [71] have questioned whether online social spaces are seen as public. They examined the information-related practices of social network users with respect to users' own information and information of others posted in online social spaces, and found that users indeed view and treat online social networks as public venues [71]. However, people feel that their privacy is violated when information about them is passed inappropriately from one context or social sphere to another [203, 451]. Sleeper et al. [432] found that users of social networks would often self-censor, or choose not to share information, but that they would have shared approximately half of the unshared content if they had been able to exactly target their desired audiences, and Liu et al. [280] showed that, although users tend to share content on social media using default privacy settings, it is usually not the intended privacy disposition. Drawing on their results, the authors suggest the possibility of correctly configuring privacy settings by leveraging the social links between

users' friends to group them into distinguished communities [280].

Additionally, research on the use of privacy settings for separating audiences withing social networks demonstrates the need for a more gradual approach to content "publicness". For example, exploration of how YouTubers manipulate access to their videos revealed that some behaviours cannot be viewed as neither strictly public nor strictly private [260]. Similarly, based on the comparison of architecture of three social network sites, Papacharissi [349] discusses the styles of self-presentation in "privately public" and "publicly private". Finally, Lankton et al. [261] discuss three privacy management strategies and refer to limiting network size as the most public strategy, whereas limiting disclosures or using privacy settings to restrict access are gradually more private [261].

Finally, the user's perception of breadth of the audience and corresponding interpretation of their space as private or public are apparent in users' concerns regarding the system's ability to "spy" on them. These concerns seem to be especially prominent in the adoption of smart homes [31] and stand-alone ubiquitous systems, e.g. home smart speakers [68, 297]. For example, Lau et al. [264] explored privacy perceptions, concerns and privacy-seeking behaviors with smart speakers and described the user's highlighted expectations of their "private home affairs to remain private", while a smart speaker can be ostensibly used to listen in on their homes.

8.3.2 Addressee

The **addressee factor** is represented by the continuum from defined to undefined addressee and refers to the degree of specificity of the expected addressee for the shared information.

While **space** and **addressee** factors are conceptually related, we argue that they represent two distinct factors. Specifically, as we mentioned, the **space factor** refers to the perceived breadth of the audience for the shared information. However, even when the audience is perceived as being broad as opposed to targeted, the specifics of this audience might still be variably defined. For example, exploring online reputation management (**project 1**, chapter 2) we identified three distinct social contexts – one's professional, social, and personal online personas. While inability to highlight content related to one of these personas and mask it for other contexts restricts the expected breadth of the audience, this audience is still rather defined, leading to the user's desire to ensure the existence and accessibility of certain content (e.g. professionally), regardless of the audience's breadth. On the other hand, an expected narrow audience might be "undefined". As an example consider communication with an unfamiliar person on the phone; the audience is

narrow and targeted, yet has a lower degree of specificity, i.e. the user has limited contextual information to form appropriate expectations on interpretation and use of shared information.

Another argument toward distinguishing between **space** and **addressee** factors is that there are different design mechanisms that can be applied towards each. For instance, in communication directly with the system, the degree of specificity of an **addressee** – how defined it is – can be approached by applying anthropomorphization-provoking design. Indeed, we find (**project 5**, chapter 6) that anthropomorphized behavioural and visual perceptions of agents consist of rich, socially-meaningful interpretations, reducing the sense of uncertainty regarding an agent as a communication partner.

Our explorations of video-mediated participation in hybrid work meetings (**project 3**, chapter 4) also demonstrates how lower degrees of contextual "definitiveness" of information recipient affect one's engagement in a communication process. For instance, lacking an ability to form contextual interpretations of remote participants' states and actions (restrictions to their signaling function) leads to their social exclusion from an active meeting process. Similarly, remote participants' lack of ability to collect contextual information required for forming appropriate "definitiveness" of the information recipient (restricted function of social information gathering) causes their lower motivation to engage.

The effects of specificity of an addressee can be seen most clearly in relation to our work on sharenting behaviours (**project 2**, chapter 3). In particular, within the intended communication context with a parent's social circle, a protagonist tends to be assigned to a parent, which amplifies the parent's desire to share the information. However, when the imagined audience becomes less defined for a sharing parent – the present and future "unintended" audience – the expected shared communication context is lost, and the interpretation of the shared information changes both for the sharing parent and for the viewers of this information. Furthermore, the defined perceptions of "unintended" audience also affect sharing patterns, e.g. limited sharing due to the expected malicious intentions revolving around socially inappropriate use of the child's photos.

However, the effects of the degree of specificity of the expected information addressee on users' willingness to share information seem to be non-linear, which suggests the underlying linkage rules between this and other contextual factors. On one hand, exploring social presence in the online learning environment, Tu and McIsaac [470] found that, due to the absence of social context cues, users tend to lose a sense of self-awareness and perceive themselves as invisible, and, therefore, are more likely to allow higher self-disclosure. On the other hand, research also shows that the degree of the addressee specificity is tightly related to the ability to form an expectation on a shared frame of reference for information

interpretation [44, 445, 433] and on rules for managing shared information [95, 355, 356]. This is reflected, for instance, in the documented strategy of bloggers and social network site users to use coded or ambiguous language to protect information they choose to share with a limited circle [92, 95].

The relationship between these expectations and addressee expectations can be considered through the lens of "familiarity" of the information recipient [408, 173], which positively affects information sharing [522, 231, 89] by reducing uncertainty [274] and supporting trust development [238, 457]. The nature of trust between the user and the receiver of information, as well as the expected commitment and reciprocity in their social relationship, seem to be major antecedents for users' willingness to share [526, 379]. At the same time, research on mediated communication discusses the effects of users' shared situational awareness [199, 190, 48, 200] and demonstrates that a communication medium's "affective capacity" [357] – how much emotional information a medium can carry, and thus, how much it can define an information sender – has direct effect on the emergence of trust in social communication situations [57, 56].

To increase the degree of addressee specificity and sense of familiarity, designers often exploit anthopomorphic features, because, as noted by Epley et al. [142], anthropomorphization serves to make sense of an unfamiliar situation. Furthermore, applied to technology, the functional role of anthropomorphization seems to affect the perceived approachibility of a system [62, 46, 285] and the level of users' trust [490, 116, 365, 267]. For example, Niu, Terken and Eggen [335] demonstrated that anthropomorphizing information may foster trust in the perception of autonomous vehicles, and Waytz, Heafner and Epley [492] found that anthropomorphic features increase users' confidence in vehicle competence [492]. Applied to AI assistants, Crone et al. [109] found that affective capabilities of virtual agents increase users' trust in the agents' affective capabilities and in their cognitive capabilities, and Yuksel et al. [524] showed that attractiveness of agents is equally or even more important for user trust in agents than agents' reliability.

8.3.3 Reason

The **sharing reason factor** refers to the primacy of the intended information perception, and stretches from "instrumental" sharing, performed as a required step to receive some service benefit, to "objective sharing" performed as the main activity. Thus, "instrumental" sharing is secondary within the main task and users' willingness for sharing is defined by the anticipated respect of the boundaries of the main task. "Objective" information sharing, on the other hand, implies that sharing is primary to the task (e.g. a post on a

social network), and the willingness for "objective" sharing is defined by the expectation on the correctness of its interpretation and corresponding social benefits.

As an example of "instrumental" sharing, consider parents' use of monitoring technologies (**project 4**, chapter 5). The child's information collected by the system (shared with the system) is perceived as highly sensitive, but parents are willing to provide this information (use the technology) if it properly addresses the desired use-cases: adjusting routine activities, performing emergency activities, and building communication and a deeper personal relationship with a child. Correspondingly, the retrieval of this information is only desired (and expected) in association with the sharing intent – when an action is required from a parent. At the same time, the expectations of "inappropriate" use of this information, e.g. for telemarketing purposes, decreases users' willingness to share.

The effects of perception of information sharing as "instrumental" can also be seen in user considerations of the resource flow in conversational agents task performance (**project 6**, chapter 7). Specifically, concerns seem to be lower for user-initiated tasks, and particularly for "assistant-user" interactions, i.e. when a conversational agent facilitates the performance of the primary activity while the user actually performs it ("help me to do" tasks). On the contrary, users express significant concerns based on their beliefs regarding an agent's unauthorized and uncontrolled access to a user's resources and unauthorized passing of these resources – information that was shared with the system for the purpose of a particular task but feared to be used outside of its scope. Note that these concerns are distinctly different from the fear of being hacked and refer specifically to user expectations on the inappropriate use of the information by the original system.

"Objective" information sharing appears, for example, in sharenting practices (**project 2**, chapter 3) and users' practices and perception of online reputation management (**project 1**, chapter 2). These projects illustrate how the expectations of social interpretations and social relevance of shared information affect users' willingness to share or to limit information dissemination. Similarly, the analysis of technology mediated participation in work meetings (**project 3**, chapter 4) further illustrates the facilitating role of expected social interpretations in willingness for "objective" information sharing. For instance, consider the limitations imposed by video-mediated meeting participation on the signaling function of attention actions and the effects of these limitations on communication engagement.

Finally, anthropomorphized perceptions of conversational agents (**project 5**, chapter 6) reveal an increasingly interesting question on the relationship between "instrumental" and "objective" information sharing within one given system. In particular, the social nature of the environment, which these systems occupy (e.g. home), in combination with the discovered richness of socially meaningful interpretations of their interactions, arguably,

encourages "objective" sharing. General-purpose conversational agents are capable of supporting diverse, purely social, non-task-based interactions (e.g. [378]). On the other hand, the main functionality of conversational agents – personal assistance to the user – implies diverse "instrumental" sharing. This observation suggests that it should be a system's responsibility to correctly recognize the acceptability and appropriateness of use of information based on the factor of sharing reason.

Multiple examples exist in the literature of the notion of differences between "instrumental" and "objective" information sharing. For example, Heckner et al. [202] compared usage patterns within social tagging systems, and found that while resource sharing emerges as an intra-system motivation, users differ with respect to social spheres of sharing. Additionally, research on remote monitoring of family members has frequently validated the utility of such systems [319, 396, 207, 290], but while elderly users value health and security aspects of monitoring ("instrumental" sharing), monitoring of activities of daily life and social connection technologies ("objective" sharing) might be met with greater resistance [423], which some research explains as balancing awareness and privacy [461]. There is also latent evidence that **sharing reason** should be considered as a continuum. Ahmed et al. [10] discovered gradual willingness of bystanders to share additional information with visually impaired users of assistive technologies – willing to share some kinds of personal information, bystanders' further sharing would require higher-security assurances by improving their control over how their information is shared.

The specific effects of different aspects of the **sharing reason** factor could be potentially best understood through the lens of the economic perspective on decision making. This approach considers privacy decisions from the perspective of cost benefit and risk reward assessment [17, 240, 355, 377]. e.g. the privacy calculus theory postulates that individuals perform a calculus between the expected loss of privacy and the potential gain of disclosure [237, 506]. Acquisti and Grossklags [5], for example, show that regardless of users' privacy concerns, they still traded-off privacy for other advantage. Dourish and Anderson [131] categorize related research approaches as *privacy as economic rationality*, where users' sharing intent is formed from a perspective of exchange value of information, so that sharing decisions are made to optimize for individual benefit and information is disclosed when a potential benefit outweighs expected losses. The differences in the effects of "objective" and "instrumental" sharing can be explained though the differences in the nature of perceived benefits they provide.

Understanding the mechanisms of "objective" sharing requires recognition of the complex relationship between privacy and sociality (e.g. [366, 487, 246, 242]). For instance, McCullagh [307] explored bloggers' subjective sense of privacy and showed that blogging offers users an important opportunity to work on their self-identity through the avail-

able degree of self-expression, bring the 'private' to the public, despite the associated privacy risks [307]. Overall, social advantages of online self-disclosure for socialization [194, 140, 458, 336, **?**], as well as social relevance of an application [458, 89] are shown to be significant predictors of users' willingness to share information.

Users' willingness to share their information "instrumentally", on the other hand, is guided by the expected balance between the service benefit and potential risk of information being used outside of the scope of the initial task for which it was provided. Rauhofer [376] refers to this as the purpose limitation principle. Correspondingly, users might attempt to falsify more sensitive personal information, e.g. when interacting with organizations such as banks, to enact privacy protection rules but to still gain the service benefit [95]. The effects of such probability assessment of information misuse depend on the perceived "sensitivity" of information – how crucial it is to be kept within the task boundaries. For example, given the "instrumental" nature of both sharing reasons, willingness to share physiological information concerning the user's health and illness [504] is reduced in comparison to the physiological data in relation to sports [475]. Further, the closer the information is connected to a user's home, the more reluctant the user becomes to share the information [264, 531, 474, 31, 68, 297]. Correspondingly, the extent of anonymization and the type of shared data become increasingly relevant attributes in "instrumental" sharing decisions, out-weighing the importance of the received benefits [530].

8.3.4 Time

The last factor of SART is the **time factor**, which refers to the perceived length of the information life-span and stretches from the perceived immediate use of information to the perceived indefinite information life-cycle. Overall, the forthcoming to the immediate end of the spectrum tends to provoke easier information sharing, while expectations of its longer life-time motivate users to limit their sharing activities. In part, this is related to the increasing chance of the information being misinterpreted (connection to the **addressee factor**), to lose its descriptive relevance over time, as well as to be used for unintended purposes (connection to the **reason factor**).

The perception of time factor and its effect on users' sharing intent can be observed in one's sharing practices with regard to their online reputation management (**project 1**, chapter 2), where the age of the information plays a significant role in the assessment of its relevance for a particular social context, and users express a desire to be able to prioritize retrieved content accordingly. Similarly, the analysis of protagonist assignment in sharenting posts (**project 2**, chapter 3) demonstrates that the "future" perceptions and

interpretations of the information are expected to shift towards the child's identity, and explores how these considerations affect parent's willingness to share the information.

The associations between the time factor and the intended instrumental reason for information sharing is also revealed through the analysis of privacy concerns in interactions with conversational agents (**project 6**, chapter 7), e.g. perceptions of agents "storing" information. Interestingly, however, it did not appear in our data on parents' information needs and use of monitoring technologies (**project 4**, chapter 5). One potential explanation for it might lay in the perceived "immediacy" of the information use and relevance due to the practical focus of use-cases on guiding routine activities.

The effects of the time factor are directly related to the hyperbolic discounting phenomena [397, 237], which refers to the bias of perceiving the benefits of information disclosure to be immediate rather than delayed, causing people to perceive risks to be lower and benefits to be higher [506]. For example, Acquisti and Grossklags empirically demonstrated that even with sufficient information to make privacy-sensitive decisions, people are likely to trade off long-term privacy for short-term benefits [6]. Hence, the hyperbolic discounting phenomena captures the differences in treatment of the present in comparison to the future. These differences become apparent when the previously "under-weighted" future finally comes but the gains from original benefits have faded, causing users to feel that information protection is not adequately supported.

These issues are further complicated by the archival nature of digital information management, which often fails to properly transfer information from one to other contexts [203, 152, 370, 296]. Although research on information retrieval recognizes the dependency of functional interest and the age of information [32, 72], the systematic understanding of complex requirements for archival management of user's information remains underdeveloped. Additionally, the prolonged information lifespan potentially affects not only the relevance, but also the meaning of the information. For example, parents commonly share information about their parenting experience which includes information about their young children [114, 98, 64, 444], but with time this information becomes a component of a child's online identity, often causing corresponding self-representation issues [114, 167].

To adapt to the resulting shortcomings of temporal context of information, users have adopted a number of strategies for protecting their privacy in light of prolonged information lifespans [203], including anonymization [530] and attempts to dis-associate information and their name (e.g. the "right to be forgotten" [220, 450, 14]). In relation to social content, Zhao et al. showed that people deliberately curate an ahistorical digital presence in an effort to keep their self-representation current [529]. Unfortunately, however, these strategies are not always successful or even accessible, leading people to feel disempowered [511].

Taken together, the appropriate support of contextual time interpretation in information management is an important aspect in the design of systems that support communication [377, 152, 32, 511].

8.3.5 Summary

In this section, we presented the SART design framework, encompassing four contextual factors affecting user's willingness to share information and their expectations on information interpretation and use. The four factors – **space, addressee, reason and time** – have been identified based on the cumulative analysis of our findings from six research project, and then investigated through the lens of existing literature. As discussed, these contextual factors are interlinked, interdependent and should not be considered in isolation. However, we suggest that the ability to draw initial distinctions between these factors in application to diverse types of systems, first, allows researchers to form a basis for efforts toward systematic and functional understanding of context, and second, allows for an informed factor-targeted approach in systems design.

Chapter 9

Future Work and Conclusion

In this thesis, we presented the results of our exploration of how context affects users' information sharing intent (their expectations on the understanding and use of information). We discussed the results from six individual research projects representing three system types, and demonstrated that when communication happens through the system the role of the system is to gather communication context unavailable to the communication endpoint and to integrate it into information communication; when it happens around the system, the system should adapt its operations to provide information in the most contextually suitable format; and when a user communicates with the system, the role of the system is to recognize and "match" this context in communication with the user. Based on cumulative analysis of our finding from six individual projects, we synthesised a set of four contextual factors affecting users' sharing intent across the system types. We presented these four factors as an early design framework, that we call SART, according to the included factors of sharing space, addressee, reason and time. In a further effort toward operationalization of these factors, we defined each of them as a continuum through a dichotomy of descriptive "values".

9.1 Future Work

While SART is an early stage framework, we see its end-goal in allowing for an informed factor-targeted approach in systems design, as well as modeling informing information sharing context. Achieving this goal requires further research in several directions.

First, the analysis of our findings allowed us to suggest an initial outline of the observed

mechanisms of how each factor influences information sharing intent. However, proper operationalization of our early SART framework requires further targeted investigation of these mechanisms. For instance, more research is needed to identify the structure (binary vs multinomial) and the nature (linear vs nonlinear) of each factor's influences. Additionally, we have discussed that the four factors, identified in this work, are interlinked, interdependent and should not be considered in isolation. Correspondingly, operationalization of the SART framework requires more research to identify relative weights of these factors and to formulate linkage (interdependency) rules of their influencing mechanisms.

Furthermore, while we argue that each factor of SART appears across system types, the specifics of their influencing mechanisms could vary between systems. An understanding of the parameters affecting these variations (e.g. an interface type or ubiquity of a system), in combination with corresponding variations in relative weights of factors and mechanisms linkage rules, would allow for a predictive use of SART by plotting a given system on framework dimensions.

Finally, while through the six individual projects we have investigated diverse communication scenarios with varying sender-receiver relationships, current work does not account for personal differences in users' information sharing preferences. Evidence from privacy research suggest that individuals differ in their privacy behaviours. Given that privacy investigations continuously demonstrate the possibility of clustering users according to their information disclosure choice-making (e.g. [93, 136, 247, 426]), understanding of the context mechanisms for each particular cluster of users yields promising opportunities for framework sensitivity.

9.2 From a Tool to a Collaborator

The system's ability to understand and act on the user's information sharing intent would not only allow to appropriately index, store, and retrieve information, but would also affect the conceptual role of the system in communication. Users' trust in a system's ability can be anticipated to shift their use-pattern from "using a tool" to "collaborating" with the system, allowing for the paradigmatic shift from "technologically extended human" toward "human-computer symbiosis" in communication [273].

Analysing research efforts in artificial intelligence and human-computer interaction, Terveen [466] suggests we view collaboration as a process "in which two or more agents work together to achieve shared goals" and distinguishes between two major approaches to the human-computer collaboration development. The first approach, to get computers

to collaborate with humans, is built on the idea of enabling computers to act "human-like" (**the human emulation approach**) [466]. The system's ability to understand users' information sharing intent, suggested in this thesis, aims at emulating human-like interpretation of users' expectations on information interpretation and handling, and thus, shares commonalities with this approach. The system's ability to understand users' sharing intent and to handle information accordingly might allow it to become a latent actor in the communication process, rather then being purely an information sharing channel, especially if considered in light of users' general tendency to treat systems as social actors (e.g., the media equation paradigm [326, 323, 324]).

However, the suggested concept of intent-grounded information handling also incorporates aspects of the second approach to human-computer collaboration, which focuses on exploiting a computer's unique abilities to complement humans (**the human complementary approach**) [466]. The distinct advantages offered by both "instrumental" [377, 530, 319, 133] and "objective" [194, 140, 458, 336, **?**] mediated information sharing are predicated upon systems' unique abilities, including time and space distributed information dissemination (e.g. teleconferencing, file sharing, social media, etc.), unique sensing abilities (e.g. physiological and location monitoring technology, etc.), and unique information processing, access and alternation abilities (e.g. search engines, recommendation systems, virtual assistants, etc.). Correspondingly, whether users interact through, around, or with the system, intent-grounded information handling exploits these systems' abilities to complement human user, e.g. to retrieve information according to the receiving context concealed from the user, to alternate communication environment, or to control information routs.

Therefore, the approach of intent-grounded information handling opens directions for advancing efforts in human-computer collaboration, combining aspects of both human emulation and human complementary approaches. Tennenhouse [465] saw the opportunity to drive human-computer symbioses [273] efforts forward by developing proactive systems and taking humans "out of the interactive loop and into a supervisory role". Twenty years later, we revisit this approach with respect to information sharing and advocate that computing should bringing the system into the loop as a latent communication actor to promote the system's role from an *information sharing channel* to a *collaborator in information sharing*.

9.3 Conclusion

In its early stages, in the 1980s, HCI as a field was predominantly concerned with productivity applications and how to make computers so that humans could understand them [81]. The following decades of expansion of interaction environments, application domains, and technical capabilities, however, allowed HCI researchers to shift their focus to how to make humans and computers understand each other.

The core research in this thesis focuses around six research projects in the area of computer-mediated communication. As noted in Figure 1.3, reproduced here as Figure 9.1, we place our six systems along a continuum based upon their role in communication, i.e., whether communication occurs through, around, or with technology. Along the continuum, the idea of contextualization during information sharing between sender and recipient (whether another person or a computing system) arises as a paramount concern. This, in turn, leads to a synthesis, in Part IV, that explores, specifically, a set of high-level contextual factors that can be used to guide systems in their treatment of information shared by a user such that information is managed in-line with users' intent. Based on cumulative analysis of our results and evidence from the literature, we suggested four contextual factors that affect users' information sharing intent across system types (SART, our early design framework) as a first step toward designing systems that more appropriately reflect users' intent during information sharing.

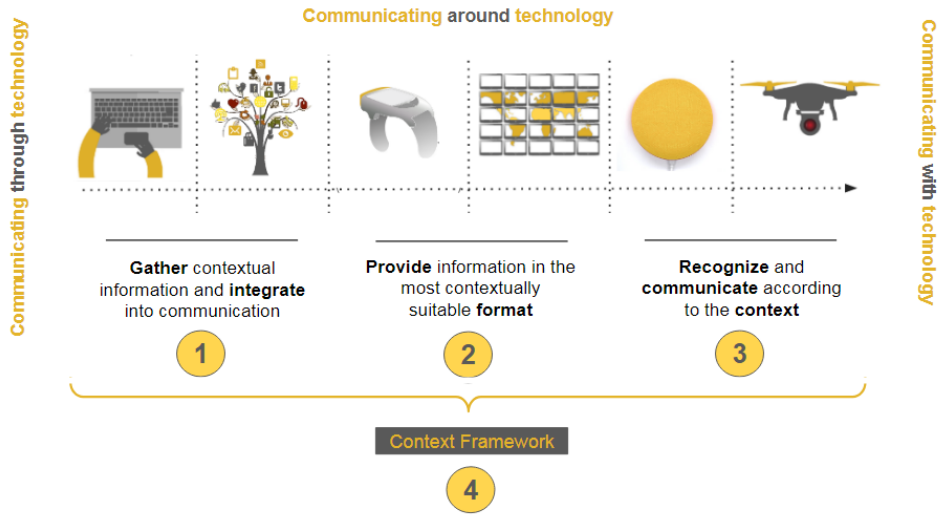

Figure 9.1: The structure of this thesis.

References

[1] Gregory D Abowd, Christopher G Atkeson, Jason Brotherton, Tommy Enqvist, Paul Gulley, and Johan LeMon. Investigating the capture, integration and access problem of ubiquitous computing in an educational setting. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 440–447, 1998.

[2] Gregory D Abowd, Christopher G Atkeson, Jason Hong, Sue Long, Rob Kooper, and Mike Pinkerton. Cyberguide: A mobile context-aware tour guide. *Wireless networks*, 3(5):421–433, 1997.

[3] Gregory D Abowd, Anind K Dey, Peter J Brown, Nigel Davies, Mark Smith, and Pete Steggles. Towards a better understanding of context and context-awareness. In *International symposium on handheld and ubiquitous computing*, pages 304–307. Springer, 1999.

[4] Alessandro Acquisti and Ralph Gross. Imagined communities: Awareness, information sharing, and privacy on the facebook. In *Privacy Enhancing Technologies*, pages 36–58. Springer Berlin Heidelberg, 2006.

[5] Alessandro Acquisti and Jens Grossklags. Privacy attitudes and privacy behavior. In *Economics of information security*, pages 165–178. Springer US, 2004.

[6] Alessandro Acquisti and Jens Grossklags. Privacy and rationality in individual decision making. *IEEE Security & Privacy*, 3(1):26–33, 2005.

[7] Anne Adams and Martina Angela Sasse. Users are not the enemy. *Communications of the ACM*, 42(12):40–46, 1999.

[8] Gediminas Adomavicius and Alexander Tuzhilin. Context-aware recommender systems. In *Recommender systems handbook*, pages 217–253. Springer, 2011.

[9] Shane Ahern, Dean Eckles, Nathaniel S Good, Simon King, Mor Naaman, and Rahul Nair. Over-exposed?: privacy patterns and considerations in online and mobile photo sharing. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 357–366. ACM, 2007.

[10] Tousif Ahmed, Apu Kapadia, Venkatesh Potluri, and Manohar Swaminathan. Up to a limit? privacy concerns of bystanders and their willingness to share additional information with visually impaired users of assistive technologies. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3):1–27, 2018.

[11] Unai Alegre, Juan Carlos Augusto, and Tony Clark. Engineering context-aware systems and applications: A survey. *Journal of Systems and Software*, 117:55–83, 2016.

[12] Irwin Altman. The environment and social behavior: Privacy, personal space, territory, and crowding. 1975.

[13] Irwin Altman. Toward a transactional perspective. In *Environment and Behavior Studies*, pages 225–255. Springer, 1990.

[14] Meg Leta Ambrose. It's about time: privacy, information life cycles, and the right to be forgotten. *Stan. Tech. L. Rev.*, 16:369, 2012.

[15] Tawfiq Ammari, Priya Kumar, Cliff Lampe, and Sarita Schoenebeck. 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*, pages 1895–1904. ACM, 2015.

[16] Janna Anderson and Lee Rainie. Millennials will make online sharing in networks a lifelong habit. *Pew Research. Retrieved from Pewinternet. org website: http://pewresearch. org/millennials*, 2010.

[17] R. Anderson. Why information security is hard – an economic perspective. In *Proceeding of 17th annual computer security applications conference, New Orleans, Louisiana*, 2001.

[18] Ofcom Annual Report. *Children and parents: Media use and attitudes report*. Ofcom, 2018.

[19] Denise Ante-Contreras. *Distracted Parenting: How Social Media Affects Parent-Child Attachment*. PhD thesis, Electronic Theses, Projects, and Dissertations, 2016.

[20] Manabu Arai, Ellen Gurman Bard, and Robin Hill. Referring and gaze alignment: Accessibility is alive and well in situated dialogue. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 31, 2009.

[21] Michael Argyle and Mark Cook. Gaze and mutual gaze. 1976.

[22] Ingrid Arreola, Zan Morris, Matthew Francisco, Kay Connelly, Kelly Caine, and Ginger White. From checking on to checking in: Designing for low socio-economic status older adults. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 1933–1936, New York, NY, USA, 2014. ACM.

[23] Google Assistant. Template actions for the google assistant, 2019.

[24] John Langshaw Austin. *How to do things with words*, volume 88. Oxford university press, 1975.

[25] Oshrat Ayalon and Eran Toch. Retrospective privacy: Managing longitudinal privacy in online social networks. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, pages 1–13, 2013.

[26] Alec Azad, Jaime Ruiz, Daniel Vogel, Mark Hancock, and Edward Lank. Territoriality and behaviour on and around large vertical publicly-shared displays. In *Proceedings of the Designing Interactive Systems Conference*, pages 468–477, 2012.

[27] Timur Bagautdinov, Alexandre Alahi, François Fleuret, Pascal Fua, and Silvio Savarese. Social scene understanding: End-to-end multi-person action localization and collective activity recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4315–4324, 2017.

[28] J Elin Bahner, Anke-Dorothea Hüper, and Dietrich Manzey. Misuse of automated decision aids: Complacency, automation bias and the impact of training experience. *International Journal of Human-Computer Studies*, 66(9):688–699, 2008.

[29] Matthias Baldauf, Schahram Dustdar, and Florian Rosenberg. A survey on context-aware systems. *International Journal of Ad Hoc and Ubiquitous Computing*, 2(4):263–277, 2007.

[30] Matthias Baldauf, Katrin Lasinger, and Peter Fröhlich. Private public screens: detached multi-user interaction with large displays through mobile augmented reality. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, pages 1–4, 2012.

[31] Nazmiye Balta-Ozkan, Rosemary Davidson, Martha Bicket, and Lorraine Whitmarsh. Social barriers to the adoption of smart homes. *Energy Policy*, 63:363–374, 2013.

[32] Ranieri Baraglia, Franco Maria Nardini, Carlos Castillo, Raffaele Perego, Debora Donato, and Fabrizio Silvestri. The effects of time on query flow graph-based models for query suggestion. In *Adaptivity, Personalization and Fusion of Heterogeneous Information*, pages 182–189. 2010.

[33] Susan B Barnes. A privacy paradox: Social networking in the united states. *First Monday*, 11(9), 2006.

[34] Dean C Barnlund. A transactional model of communication. In *Communication theory*, pages 47–57. Routledge, 2017.

[35] Lawrence W Barsalou. Context-independent and context-dependent information in concepts. *Memory & cognition*, 10(1):82–93, 1982.

[36] et al. Bartholomew, Mitchell K. New parents‘ facebook use at the transition to parenthood. *Family relations*, 61(3):455–469, 2012.

[37] Andy Beal and Judy Strauss. *Radically transparent: Monitoring and managing reputations online*. John Wiley & Sons, 2009.

[38] Punam Bedi and Chhavi Sharma. Community detection in social networks. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 6(3):115–135, 2016.

[39] Diana Beirl, Nicola Yuill, and Yvonne Rogers. Using voice assistant skills in family life. 2019.

[40] Ardion D Beldad and Sabrina M Hegner. More photos from me to thee: Factors influencing the intention to continue sharing personal photos on an online social networking (osn) site among young adults in the netherlands. *International Journal of Human–Computer Interaction*, 33(5):410–422, 2017.

[41] Victoria Bellotti and Keith Edwards. Intelligibility and accountability: human considerations in context-aware systems. *Human–Computer Interaction*, 16(2-4):193–212, 2001.

[42] Richard Berendsen, Bogomil Kovachev, Evangelia-Paraskevi Nastou, Maarten De Rijke, and Wouter Weerkamp. Result disambiguation in web people search. In *European Conference on Information Retrieval*, pages 146–157. Springer, 2012.

[43] Bettina Berendt, Oliver Günther, and Sarah Spiekermann. Privacy in e-commerce: Stated preferences vs. actual behavior. *Communications*, 48(4):101–106, 2005.

[44] David K Berlo. Communication as process: Review and commentary. *Annals of the International Communication Association*, 1(1):11–27, 1977.

[45] Shona M Bettany, Shona M Bettany, Ben Kerrane, and Ben Kerrane. The sociomateriality of parental style: negotiating the multiple affordances of parenting and child welfare within the new child surveillance technology market. *European Journal of Marketing*, 50(11):2041–2066, 2016.

[46] Timothy Bickmore and Justine Cassell. Social dialongue with embodied conversational agents. In *Advances in natural multimodal dialogue systems*, pages 23–54. Springer, 2005.

[47] Frank Biocca and Chad Harms. Defining and measuring social presence: Contribution to the networked minds theory and measure. *Proceedings of PRESENCE*, 2002:7–36, 2002.

[48] Frank Biocca, Chad Harms, and Jenn Gregg. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *4th annual international workshop on presence, Philadelphia, PA*, pages 1–9, 2001.

[49] Elina Birmingham, Walter F Bischof, and Alan Kingstone. Gaze selection in complex social scenes. *Visual Cognition*, 16(2-3):341–355, 2008.

[50] Elina Birmingham, Walter F Bischof, and Alan Kingstone. Social attention and real-world scenes: The roles of action, competition and social content. *The Quarterly Journal of Experimental Psychology*, 61(7):986–998, 2008.

[51] Jeremy Birnholtz, Lindsay Reynolds, Eli Luxenberg, Carl Gutwin, and Maryam Mustafa. Awareness beyond the desktop: exploring attention and distraction with a projected peripheral-vision display. In *Proceedings of Graphics Interface 2010*, pages 55–62. Canadian Information Processing Society, 2010.

[52] Milan Z Bjelica, Bojan Mrazovac, Istvan Papp, and Nikola Teslic. Context-aware platform with user availability estimation and light-based announcements. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 43(5):1228–1239, 2013.

[53] Ben Bland. Alexa, hug me: Exploring human-machine emotional relations, 1 2018. Retrieved August 21, 2019 from https://medium.com/@ben.bland/alexa-hug-me-exploring-human-machine-emotional=-relations-1f0f6e04e1db.

[54] Jim Blascovich, Jack Loomis, Andrew C Beall, Kimberly R Swinth, Crystal L Hoyt, and Jeremy N Bailenson. Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 13(2):103–124, 2002.

[55] Alicia Blum-Ross and Sonia Livingstone. "sharenting," parent blogging, and the boundaries of the digital self. *Popular Communication*, 15(2):110–125, 2017.

[56] Nathan Bos, Darren Gergle, Judith S Olson, and Gary M Olson. Being there versus seeing there: Trust via video. In *CHI'01 extended abstracts on human factors in computing systems*, pages 291–292, 2001.

[57] Nathan Bos, Judy Olson, Darren Gergle, Gary Olson, and Zach Wright. Effects of four computer-mediated communications channels on trust development. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 135–140, 2002.

[58] Lauren Bradel, Alex Endert, Kristen Koch, Christopher Andrews, and Chris North. Large high resolution displays for co-located collaborative sensemaking: Display usage and territoriality. *International Journal of Human-Computer Studies*, 71(11):1078–1088, 2013.

[59] Sheryl Brahnam and Antonella De Angeli. Gender affordances of conversational agents. *Interacting with Computers*, 24(3):139–153, 2012.

[60] Sheryl Brahnam and Margaret Weaver. Re/framing virtual conversational partners: A feminist critique and tentative move towards a new design paradigm. In *International Conference of Design, User Experience, and Usability*, pages 172–183. Springer, 2015.

[61] Eva Brandt. Designing exploratory design games: a framework for participation in participatory design? In *Proceedings of the ninth conference on Participatory design: Expanding boundaries in design-Volume 1*, pages 57–66, 2006.

[62] Cynthia Breazeal. Toward sociable robots. *Robotics and Autonomous Systems*, 42(3):167 – 175, 2003. Socially Interactive Robots.

[63] Ingar Brinck. Attention and the evolution of intentional communication. *Pragmatics & Cognition*, 9(2):259–277, 2001.

[64] Anna Brosch. When the child is born into the internet: Sharenting as a growing trend among parents on facebook. 2016.

[65] Nela Brown and Tony Stockman. Examining the use of thematic analysis as a tool for informing design of new family communication technologies. In *Proceedings of the 27th International BCS Human Computer Interaction Conference*, BCS-HCI '13, pages 21:1–21:6, Swinton, UK, UK, 2013. British Computer Society.

[66] Peter J Brown, John D Bovey, and Xian Chen. Context-aware applications: from the laboratory to the marketplace. *IEEE personal communications*, 4(5):58–64, 1997.

[67] Allison Bruce, Illah Nourbakhsh, and Reid Simmons. The role of expressiveness and attention in human-robot interaction. In *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, volume 4, pages 4138–4142. IEEE, 2002.

[68] Raluca Budiu and Page Laubheimer. Intelligent assistants: Creepy, childish, or a tool? users' attitudes toward alexa, google assistant, and siri, 8 2018. Retrieved February 22, 2019 from https://www.nngroup.com/articles/voice-assistant-attitudes/?lm=intelligent-assistants-poor-usability-high-adoption&pt=article.

[69] Raluca Budiu and Page Laubheimer. Intelligent assistants have poor usability: A user study of alexa, google assistant, and siri, 7 2018. Retrieved February 21, 2019 from https://www.nngroup.com/articles/intelligent-assistant-usability/.

[70] Raluca Budiu and Kathryn Whitenton. The paradox of intelligent assistants: Poor usability, high adoption, 9 2018. Retrieved February 21, 2019 from https://www.nngroup.com/articles/intelligent-assistants-poor-usability-high-adoption/.

[71] Jacquelyn Burkell, Alexandre Fortier, Lorraine (Lola) Yeung Cheryl Wong, and Jennifer Lynn Simpson. Facebook: Public space, or private space? *Information, Communication & Society*, 17(8):974–985, 2014.

[72] Gary Burnett. Information worlds and interpretive practices: Toward an integration of domains. *Journal of Information Science Theory and Practice*, 3(3):6–16, 2015.

[73] Bill Buxton. *Sketching user experiences: getting the design right and the right design*. Morgan kaufmann, 2010.

[74] Guoray Cai and Yinkun Xue. Activity-oriented context-aware adaptation assisting mobile geo-spatial activities. In *Proceedings of the 11th international conference on Intelligent user interfaces*, pages 354–356, 2006.

[75] Jose Camacho-Collados and Mohammad Taher Pilehvar. From word to sense embeddings: A survey on vector representations of meaning. *Journal of Artificial Intelligence Research*, 63:743–788, 2018.

[76] Jean Camp and Kay Connelly. Beyond consent: privacy in ubiquitous computing (ubicomp). *Digital privacy: Theory, technologies, and practices*, pages 327–343, 2008.

[77] Statistics Canada. Education indicators in canada: An international perspective, 2018.

[78] Nil Goksel Canbek and Mehmet Emin Mutlu. On the track of artificial intelligence: Learning with intelligent personal assistants. *Journal of Human Sciences*, 13(1):592–601, 2016.

[79] Han Cao et al. Crossmodal attention in public-private displays. In *2006 ACS/IEEE international conference on pervasive services*, pages 13–18. IEEE, 2006.

[80] Mihaela Cardei, Brandon Jones, and Daniel Raviv. A pattern for context-aware navigation. In *20th Conference on Pattern Languages of Programs (PLoP). PLoP*, volume 13, pages 23–26. Citeseer, 2013.

[81] John M Carroll. Human computer interaction-brief intro. *The Encyclopedia of Human-Computer Interaction, 2nd Ed.*, 2013.

[82] Justine Cassell. More than just another pretty face: Embodied conversational interface agents. *Communications of the ACM*, 43(4):70–78, 2000.

[83] Richard Catrambone. Anthropomorphic agents as a user interface paradigm : Experimental findings and a framework for research. In *Technical Report GIT-GVU-02-10*. Georgia Institute of Technology, 2002.

[84] Richard Catrambone, John Stasko, and Jun Xiao. Anthropomorphic agents as a user interface paradigm: Experimental findings and a framework for research. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 24, 2002.

[85] Pew Research Center. Nearly half of americans use digital voice assistants, mostly on their smartphones, 12 2017. Retrieved February 28, 2019 from https://www.pewresearch.org/fact-tank/2017/12/12/nearly-half-of-americans-use-digital-voice-assistants-mostly-on-their-smartphones/.

[86] Dave Chaffey. *Global social media research summary*. Ofcom, February 2019.

[87] Charlotte Chalklen and Heather Anderson. Mothering on facebook: exploring the privacy/openness paradox. *Social Media+ Society*, 3(2):2056305117707187, 2017.

[88] Kathy Charmaz. *Constructing grounded theory*. Sage, 2014.

[89] Xi Chen, MingXue Sun, Dong Wu, and Xiao Yu Song. Information-sharing behavior on wechat moments: The role of anonymity, familiarity, and intrinsic motivation. *Frontiers in Psychology*, 10:2540, 2019.

[90] N. Chesley. Blurring boundaries? Linking technology use, spillover, individual distress and family satisfaction. *Journal of Marriage and Family*, 67(5):1237–1248, 2005.

[91] Ming Cheung and James She. Detecting social signals in user-shared images for connection discovery using deep learning. *IEEE Transactions on Multimedia*, 2019.

[92] Jeffrey T Child and Esther A Agyeman-Budu. Blogging privacy management rule development: The impact of self-monitoring skills, concern for appropriateness, and blogging frequency. *Computers in Human Behavior*, 26(5):957–963, 2010.

[93] Jeffrey T Child, Paul M Haridakis, and Sandra Petronio. Blogging privacy rule orientations, privacy management, and content deletion practices: The variability of online privacy management activity at different stages of social media use. *Computers in Human Behavior*, 28(5):1859–1872, 2012.

[94] Jeffrey T Child, Judy C Pearson, and Sandra Petronio. Blogging, communication, and privacy management: Development of the blogging privacy management measure. *Journal of the American Society for Information Science and Technology*, 60(10):2079–2094, 2009.

[95] Jeffrey T Child and Sandra Petronio. Unpacking the paradoxes of privacy in cmc relationships: The challenges of blogging and relational communication on the internet. *Computer-mediated communication in personal relationships*, pages 21–40, 2011.

[96] Erika Chin, Adrienne Porter Felt, Vyas Sekar, and David Wagner. Measuring user confidence in smartphone security and privacy. In *Proceedings of the eighth symposium on usable privacy and security*, pages 1–16, 2012.

[97] Minji Cho, Sang-su Lee, and Kun-Pyo Lee. 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*, pages 1557–1569, 2019.

[98] Grace Yiseul Choi and Jennifer Lewallen. "say instagram, kids!": Examining sharenting and children's digital representations on instagram. *Howard Journal of Communications*, pages 1–21, 2017.

[99] Hyunji Chung, Michaela Iorga, Jeffrey Voas, and Sangjin Lee. "alexa, can i trust you?". *Computer*, 50:100–104, 09 2017.

[100] Herbert H Clark and Deanna Wilkes-Gibbs. Referring as a collaborative process. *Cognition*, 22(1):1–39, 1986.

[101] L.S. Clark. *The parent app: Understanding families in the digital age*. Oxford: Oxford University Press, 2013.

[102] Jessica L Collett. What kind of mother am i? impression management and the social construction of motherhood. *Symbolic Interaction*, 28(3):327–347, 2005.

[103] Sunny Consolvo, Ian E Smith, Tara Matthews, Anthony LaMarca, Jason Tabert, and Pauline Powledge. Location disclosure to social relations: why, when, & what people want to share. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 81–90, 2005.

[104] Reputation Management Consultants. Accessed: 2020-06-02.

[105] Gregory Conti. Could googling take down a president? *Communications of the ACM*, 51(1):71–73, 2008.

[106] Sobiesk E. Conti, G. An honest man has nothing to fear. In *Proc. SOUPS 2007*, pages 112–121. ACM Press, 2007.

[107] Juliet M Corbin and Anselm Strauss. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology*, 13(1):3–21, 1990.

[108] Joëlle Coutaz, James L Crowley, Simon Dobson, and David Garlan. Context is key. *Communications of the ACM*, 48(3):49–53, 2005.

[109] Tim Crone, Jonas Valbjørn Andersen, Mahmood Shafeie Zargar, and Bart van den Hooff. Trust me i am human: How the anthropomorphization of ai is changing the human-algorithmic relationship. 2019.

[110] Guoqiang Cui, Barbara Lockee, and Cuiqing Meng. Building modern online social presence: A review of social presence theory and its instructional design implications for future trends. *Education and information technologies*, 18(4):661–685, 2013.

[111] C. S. Russell D. H. Olson and D. H. Sprenkle. Circumplex model of marital and family systems: Vl. Theoretical update. *Family process*, 22(1):69–83, 1983.

[112] Richard L Daft and Robert H Lengel. Organizational information requirements, media richness and structural design. *Management science*, 32(5):554–571, 1986.

[113] Russell Eliot Dale. *The theory of meaning*. City University of New York, 1996.

[114] Nadine Davidson-Wall. "mum, seriously!": Sharenting the new social trend with no opt-out. In *Paper presented at the 9th Debating Communities and Social Networks 2018 OUA Conference. Online conference*, 2018.

[115] Jenny L Davis and Nathan Jurgenson. Context collapse: Theorizing context collusions and collisions. *Information, communication & society*, 17(4):476–485, 2014.

[116] Ewart J de Visser, Samuel S Monfort, Ryan McKendrick, Melissa AB Smith, Patrick E McKnight, Frank Krueger, and Raja Parasuraman. Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, 22(3):331, 2016.

[117] Madeleine Dean. Skype now supports speaker view in group calls, 3 2019. Retrieved September 07, 2019 from https://windowsreport.com/skype-speaker-view-group-calls/.

[118] Riley J. Dennis. What are the best child tracking tools and apps?, 2016.

[119] Paco Developers. The paco application, 2016.

[120] Anind K Dey. Context-aware computing: The cyberdesk project. In *Proceedings of the AAAI 1998 Spring Symposium on Intelligent Environments*, pages 51–54, 1998.

[121] Anind K Dey. Understanding and using context. *Personal and ubiquitous computing*, 5(1):4–7, 2001.

[122] Anind K Dey, Gregory D Abowd, and Daniel Salber. A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human–Computer Interaction*, 16(2-4):97–166, 2001.

[123] Anind K. Dey and Ed de Guzman. From awareness to connectedness: the design and deployment of presence displays. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2006.

[124] Joan Morris DiMicco and David R Millen. Identity management: multiple presentations of self in facebook. In *Proceedings of the 2007 international ACM conference on Supporting group work*, pages 383–386, 2007.

[125] Michael Dimock. Defining generations: Where millennials end and post-millennials begin. https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/, 2019.

[126] Lei Ding and Alper Yilmaz. Inferring social relations from visual concepts. In *2011 International Conference on Computer Vision*, pages 699–706. IEEE, 2011.

[127] Gwyneth Doherty-Sneddon and Fiona G Phelps. Gaze aversion: A response to cognitive or social difficulty? *Memory & cognition*, 33(4):727–733, 2005.

[128] Judith Donath and Danah Boyd. Public displays of connection. *bt technology Journal*, 22(4):71–82, 2004.

[129] Jay Donovan. *The average age for a child getting their first smartphone is now 10.3 years*. TechCrunch, 2016.

[130] Paul Dourish. What we talk about when we talk about context. *Personal and ubiquitous computing*, 8(1):19–30, 2004.

[131] Paul Dourish and Ken Anderson. Collective information practice: Exploring privacy and security as social and cultural phenomena. *Human-computer interaction*, 21(3):319–342, 2006.

[132] John V Draper, David B Kaber, and John M Usher. Telepresence. *Human factors*, 40(3):354–375, 1998.

[133] Lynsey Dubbeld. Telemonitoring of cardiac patients: User-centred research as input for surveillance theories. *Theorizing surveillance: The panopticon and beyond*, pages 182–205, 2006.

[134] Brian R Duffy. Anthropomorphism and the social robot. *Robotics and autonomous systems*, 42(3-4):177–190, 2003.

[135] Brian R Duffy. Anthropomorphism and the social robot. *Robotics and autonomous systems*, 42(3-4):177–190, 2003.

[136] Janna Lynn Dupree, Richard Devries, Daniel M Berry, and Edward Lank. Privacy personas: Clustering users via attitudes and behaviors toward security practices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5228–5239, 2016.

[137] Mary T Dzindolet, Scott A Peterson, Regina A Pomranky, Linda G Pierce, and Hall P Beck. The role of trust in automation reliance. *International journal of human-computer studies*, 58(6):697–718, 2003.

[138] Aarthi Easwara Moorthy and Kim-Phuong L Vu. Privacy concerns for use of voice activated personal assistant in the public space. *International Journal of Human-Computer Interaction*, 31(4):307–335, 2015.

[139] Jacob Eisenstein, Jean Vanderdonckt, and Angel Puerta. Adapting to mobile contexts with user-interface modeling. In *Proceedings Third IEEE Workshop on Mobile Computing Systems and Applications*, pages 83–92. IEEE, 2000.

[140] Nicole B Ellison, Jessica Vitak, Charles Steinfield, Rebecca Gray, and Cliff Lampe. Negotiating privacy concerns and social capital needs in a social media environment. In *Privacy online*, pages 19–32. Springer, 2011.

[141] Mica R Endsley. Measurement of situation awareness in dynamic systems. *Human factors*, 37(1):65–84, 1995.

[142] N. Epley, A. Waytz, and J. T. Cacioppo. On seeing human: A three-factor theory of anthropomorphism. In *Psychological Review*, volume 114, pages 864–886, 2007.

[143] Nicholas Epley, Scott Akalis, Adam Waytz, and John T Cacioppo. Creating social connection through inferential reproduction: Loneliness and perceived agency in gadgets, gods, and greyhounds. *Psychological science*, 19(2):114–120, 2008.

[144] Florian Eyben, Felix Weninger, Lucas Paletta, and Björn W Schuller. The acoustics of eye contact: detecting visual attention from conversational audio cues. In *Proceedings of the 6th workshop on Eye gaze in intelligent human machine interaction: gaze in multimodal interaction*, pages 7–12. ACM, 2013.

[145] F. Eyssel, L. de Ruiter, D. Kuchenbrandt, S. Bobinger, and F. Hegel. 'if you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. In *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 125–126, March 2012.

[146] Friederike Eyssel, Laura De Ruiter, Dieta Kuchenbrandt, Simon Bobinger, and Frank Hegel. 'if you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. In *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 125–126. IEEE, 2012.

[147] Friederike Eyssel and Frank Hegel. (s) he's got the look: Gender stereotyping of robots 1. *Journal of Applied Social Psychology*, 42(9):2213–2230, 2012.

[148] Hart Research Associates Family Online Safety Institute. Parents, privacy & technology use. https://www.fosi.org/policy-research/parents-privacy-technology-use/, 2015. Accessed: 2018-03-12.

[149] Xiaocong Fan, Michael McNeese, Bingjun Sun, Timothy Hanratty, Laurel Allender, and John Yen. Human–agent collaboration for time-stressed multicontext decision making. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(2):306–320, 2009.

[150] Sidney Fels, Yasuyuki Sumi, Tameyuki Etani, Nicolas Simonet, K Kobayshi, and Kenji Mase. Progress of c-map: A context-aware mobile assistant. In *Proceedings of AAAI 1998 Spring Symposium on Intelligent Environments*, pages 60–67, 1998.

[151] Julia Fink. Anthropomorphism and human likeness in the design of robots and human-robot interaction. In *International Conference on Social Robotics*, pages 199–208. Springer, 2012.

[152] Lev Finkelstein, Evgeniy Gabrilovich, Yossi Matias, Ehud Rivlin, Zach Solan, Gadi Wolfman, and Eytan Ruppin. Placing search in context: The concept revisited. In *Proceedings of the 10th international conference on World Wide Web*, pages 406–414, 2001.

[153] Kathleen E Finn, Abigail J Sellen, and Sylvia B Wilbur. *Video-mediated communication*. L. Erlbaum Associates Inc., 1997.

[154] John Fiske. *Introduction to communication studies*. Routledge, 2010.

[155] Rosemary Flanagan and Robert W Motta. Figure drawings: A popular method. *Psychology in the Schools*, 44(3):257–270, 2007.

[156] Dinei Florencio and Cormac Herley. A large-scale study of web password habits. In *Proceedings of the 16th international conference on World Wide Web*, pages 657–666, 2007.

[157] Luciano Floridi. *Information: A very short introduction*. OUP Oxford, 2010.

[158] James Fogarty, Jennifer Lai, and Jim Christensen. Presence versus availability: the design and evaluation of a context-aware communication client. *International Journal of Human-Computer Studies*, 61(3):299–317, 2004.

[159] Sean Follmer, Hayes Raffle, Janet Go, Rafael Ballagas, and Hiroshi Ishii. Video play: Playful interactions in video conferencing for long-distance families with young children. In *Proceedings of the 9th International Conference on Interaction Design and Children*, IDC '10, pages 49–58, New York, NY, USA, 2010. ACM.

[160] Adam Fourney, Ryen W. White, and Eric Horvitz. Exploring time-dependent concerns about pregnancy and childbirth from search logs. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 737–746, New York, NY, USA, 2015. ACM.

[161] Megan Freeth, Tom Foulsham, and Alan Kingstone. What affects social attention? social presence, eye contact and autistic traits. *PloS one*, 8(1):e53286, 2013.

[162] Alexandra Frischen, Andrew P Bayliss, and Steven P Tipper. Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological bulletin*, 133(4):694, 2007.

[163] David Frohlich, Allan Kuchinsky, Celine Pering, Abbe Don, and Steven Ariss. Requirements for photoware. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, pages 166–175. ACM, 2002.

[164] David Frohlich, Allan Kuchinsky, Celine Pering, Abbe Don, and Steven Ariss. Requirements for photoware. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, pages 166–175. ACM, 2002.

[165] Samratul Fuady, Masato Orishige, Haoyan Li, Hironori Mitake, and Shoichi Hasegawa. Natural interaction in asymmetric teleconference using stuffed-toy avatar robot. In *ICAT-EGVE*, pages 93–98, 2016.

[166] Eoghan Furey and Juanita Blue. She knows too much – voice command devices and privacy. In *2018 29th Irish Signals and Systems Conference (ISSC)*, pages 1–6, June 2018.

[167] Ouvrein Gaëlle and Verswijvel Karen. Sharenting: Parental adoration or public humiliation? a focus group study on adolescents' experiences with sharenting against the background of their own impression management. *Children and Youth Services Review*, 2019.

[168] Susan Gal. A semiotics of the public/private distinction. *Differences: a journal of feminist cultural studies*, 13(1):77–95, 2002.

[169] Y. Gao, Z. Pan, H. Wang, and G. Chen. Alexa, my love: Analyzing reviews of amazon echo. In *2018 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pages 372–380, Oct 2018.

[170] Yang Gao, Zhengyu Pan, Honghao Wang, and Guanling Chen. Alexa, my love: Analyzing reviews of amazon echo. In *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pages 372–380. IEEE, 2018.

[171] Albert Gatt and Emiel Krahmer. Survey of the state of the art in natural language generation: Core tasks, applications and evaluation. *Journal of Artificial Intelligence Research*, 61:65–170, 2018.

[172] William W Gaver, Abigail Sellen, Christian Heath, and Paul Luff. One is not enough: Multiple views in a media space. In *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems*, pages 335–341. ACM, 1993.

[173] David Gefen. E-commerce: the role of familiarity and trust. *Omega*, 28(6):725–737, 2000.

[174] Darren Gergle, Robert E Kraut, and Susan R Fussell. Using visual information for grounding and awareness in collaborative tasks. *Human–Computer Interaction*, 28(1):1–39, 2013.

[175] Lorna Gibson and Vicki L. Hanson. Digital motherhood: How does technology help new mothers? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, 2013.

[176] Anthony Giddens. *The constitution of society: Outline of the theory of structuration*. Univ of California Press, 1984.

[177] Arthur M Glenberg, Jennifer L Schroeder, and David A Robertson. Averting the gaze disengages the environment and facilitates remembering. *Memory & cognition*, 26(4):651–658, 1998.

[178] Ulrich Gnewuch, Stefan Morana, and Alexander Maedche. Towards designing cooperative and social conversational agents for customer service. In *ICIS*, 2017.

[179] Charles Goodwin. Action and embodiment within situated human interaction. *Journal of pragmatics*, 32(10):1489–1522, 2000.

[180] Charles Goodwin. Pointing as situated practice. In *Pointing*, pages 225–250. Psychology Press, 2003.

[181] David M Grayson and Andrew F Monk. Are you looking at me? eye contact and desktop video conferencing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 10(3):221–243, 2003.

[182] Martin Grayson, Anja Thieme, Rita Marques, Daniela Massiceti, Ed Cutrell, and Cecily Morrison. A dynamic ai system for extending the capabilities of blind people. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–4, 2020.

[183] Michael D Greenberg, Matthew W Easterday, and Elizabeth M Gerber. Critiki: A scaffolded approach to gathering design feedback from paid crowdworkers. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, pages 235–244, 2015.

[184] Saul Greenberg and Bill Buxton. Usability evaluation considered harmful (some of the time). In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 111–120, 2008.

[185] Saul Greenberg and Hideaki Kuzuoka. Using digital but physical surrogates to mediate awareness, communication and privacy in media spaces. *Personal Technologies*, 3(4):182–198, 1999.

[186] Herbert P Grice. Logic and conversation. In *Speech acts*, pages 41–58. Brill, 1975.

[187] Ralph Gross and Alessandro Acquisti. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pages 71–80. ACM, 2005.

[188] Jonathan Grudin. Desituating action: Digital representation of context. *Human–Computer Interaction*, 16(2-4):269–286, 2001.

[189] Tamy Guberek, Allison McDonald, Sylvia Simioni, Abraham H Mhaidli, Kentaro Toyama, and Florian Schaub. Keeping a low profile?: Technology, risk and privacy among undocumented immigrants. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 114. ACM, 2018.

[190] Carl Gutwin and Saul Greenberg. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)*, 11(3-4):411–446, 2002.

[191] Ido Guy, Sigalit Ur, Inbal Ronen, Sara Weber, and Tolga Oral. Best faces forward: a large-scale study of people search in the enterprise. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1775–1784, 2012.

[192] Jacek Gwizdka. What's in the context. In *Computer Human Interaction*, volume 2000, pages 1–4, 2000.

[193] Joy E Hanna and Susan E Brennan. Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language*, 57(4):596–615, 2007.

[194] Eszter Hargittai and Alice Marwick. "what can i really do?" explaining the privacy paradox with online apathy. *International Journal of Communication*, 10:21, 2016.

[195] E. Harmon and M. Mazmanian. Stories of the smartphone in everyday discourse: Conflict, tension and instability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1051–1060, 2013.

[196] Steve Harrison. A brief history of media space research and mediated life. In *Media Space 20+ Years of Mediated Life*, pages 9–16. Springer, 2009.

[197] Khalid Haruna, Maizatul Akmar Ismail, Suhendroyono Suhendroyono, Damiasih Damiasih, Adi Cilik Pierewan, Haruna Chiroma, and Tutut Herawan. Context-aware

recommender system: A review of recent developmental process and future research direction. *Applied Sciences*, 7(12):1211, 2017.

[198] Nick Haslam. Dehumanization: An integrative review. *Personality and social psychology review*, 10(3):252–264, 2006.

[199] Christian Heath and Paul Luff. Disembodied conduct: communication through video in a multi-media office environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 99–103. ACM, 1991.

[200] Christian Heath, Marcus Sanchez Svensson, Jon Hindmarsh, Paul Luff, and Dirk Vom Lehn. Configuring awareness. *Computer Supported Cooperative Work (CSCW)*, 11(3-4):317–347, 2002.

[201] Douglas D Heckathorn. Respondent-driven sampling: a new approach to the study of hidden populations. *Social problems*, 44(2):174–199, 1997.

[202] Markus Heckner, Michael Heilemann, and Christian Wolff. Personal information management vs. resource sharing: Towards a model of information behavior in social tagging systems. In *Third International AAAI Conference on Weblogs and Social Media*, 2009.

[203] Pekka Henttonen. Privacy as an archival problem and a solution. *Archival Science*, 17(3):285–303, 2017.

[204] K. Hertlein. Digital dwelling: Technology in couple and family relationships. *Family Relations*, 61(3):374–387, 2012.

[205] Sobel K. Suh H. Sung Y. C. Lee C. P. Kientz-J. A. Hiniker, A. Texting while parenting: How adults use mobile phones while caring for children at the playground. In *Proceedings of the ACM conference on human factors in computing systems*, pages 727–736, 2015.

[206] Birger Hjørland. Information: Objective or subjective/situational? *Journal of the American society for information science and technology*, 58(10):1448–1456, 2007.

[207] Sebastian Hoberg, Ludger Schmidt, Axel Hoffmann, Matthias Söllner, Jan Marco Leimeister, Christian Voigtmann, Klaus David, Julia Zirfas, and Alexander Roßnagel. Socially acceptable design of a ubiquitous system for monitoring elderly family members. 2012.

[208] Thomas Hofer, Wieland Schwinger, Mario Pichler, Gerhard Leonhartsberger, Josef Altmann, and Werner Retschitzegger. Context-awareness on mobile devices-the hydrogen approach. In *36th annual Hawaii international conference on system sciences, 2003. Proceedings of the*, pages 10–pp. IEEE, 2003.

[209] S. L. Hofferth and J. Sandberg. How American children spend their time. *Journal of Marriageand the Family*, 63:295–308, 2001.

[210] Jim Hollan and Scott Stornetta. Beyond being there. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 119–125. ACM, 1992.

[211] Eric Horvitz, Edward Cutrell, and Mary Czerwinski. Notification, disruption, and memory: Effects of messaging interruptions on memory and performance. In *Proceedings of INTERACT*, page 263, 2001.

[212] Eric Horvitz, Carl Kadie, Tim Paek, and David Hovel. Models of attention in computing and communication: from principles to applications. *Communications of the ACM*, 46(3):52–59, 2003.

[213] Eric Horvitz and Tim Paek. Harnessing models of users' goals to mediate clarification dialog in spoken language systems. In *International Conference on User Modeling*, pages 3–13. Springer, 2001.

[214] Bernardo A Huberman and Fang Wu. The economics of attention: maximizing user value in information-rich environments. In *Proceedings of the 1st international workshop on Data mining and audience intelligence for advertising*, pages 16–20. ACM, 2007.

[215] R. Hughes and J. Hans. Computers, the internet and families: A review of the role new technology plays in family life. *Journal of Family Issues*, 22(6):776–790, 2001.

[216] Giulio Iacucci and Kari Kuutti. Everyday life as a stage in creating and performing scenarios for wireless devices. *Personal and Ubiquitous Computing*, 6(4):299–306, 2002.

[217] Judith Irvine, Bambi Schieffelin, CM Series, Marjorie Harness Goodwin, Joel Kuipers, Don Kulick, John Lucy, Elinor Ochs, et al. *Rethinking context: Language as an interactive phenomenon*. Number 11. Cambridge University Press, 1992.

[218] Catherine Jackson and Angela Orebaugh. A study of security and privacy issues associated with the amazon echo. *International Journal of Internet of Things and Cyber-Assurance*, 1(1):91–100, 2018.

[219] Mikkel R Jakobsen and Kasper Hornbæk. Negotiating for space?: Collaborative work using a wall display with mouse and touch input. In *CHI*, pages 2050–2061, 2016.

[220] Meg Leta Jones. *Ctrl+ Z: The right to be forgotten*. NYU Press, 2018.

[221] Rodney H Jones. 14 sites of engagement as sites of attention: Time, space and culture in electronic discourse. 2005.

[222] Sanjay Kairam, Mike Brzozowski, David Huffaker, and Ed Chi. Talking in circles: selective sharing in google+. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1065–1074, 2012.

[223] Victor Kaptelinin and Bonnie A Nardi. *Acting with technology: Activity theory and interaction design*. MIT press, 2006.

[224] Kari Daniel Karjalainen, Anna Elisabeth Sofia Romell, Photchara Ratsamee, Asim Evren Yantac, Morten Fjeld, and Mohammad Obaid. Social drone companion for the home environment: A user-centric exploration. In *Proceedings of the 5th International Conference on Human Agent Interaction*, pages 89–96. ACM, 2017.

[225] Bahareh Ebadifar Keith and Stacey Steinberg. Parental sharing on the internet: Child privacy in the age of social media and the pediatrician's role. *JAMA pediatrics*, 171(5):413–414, 2017.

[226] Alice Kerry, Richard Ellis, and Susan Bull. Conversational agents in e-learning. In *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pages 169–182. Springer, 2008.

[227] Jan H Kietzmann, Kristopher Hermkens, Ian P McCarthy, and Bruno S Silvestre. Social media? get serious! understanding the functional building blocks of social media. *Business horizons*, 54(3):241–251, 2011.

[228] Cory A. Kildare and Wendy Middlemiss. Impact of parents mobile device use on parent-child interaction: A literature review. *Computers in Human Behavior*, 2017.

[229] Aelee Kim, Jooyun Han, Younbo Jung, and Kwanmin Lee. The effects of familiarity and robot gesture on user acceptance of information. In *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on*, pages 159–160. IEEE, 2013.

[230] Bomyeong Kim, Hyun Young Kim, and Jinwoo Kim. Getting home safely with drone. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 117–120. ACM, 2016.

[231] Hyang-Sook Kim. What drives you to check in on facebook? motivations, privacy concerns, and mobile phone involvement for location-based information sharing. *Computers in Human Behavior*, 54:397–406, 2016.

[232] Bret Kinsella. What people ask their smart speakers, 2018.

[233] Julia Kiseleva, Kyle Williams, Jiepu Jiang, Ahmed Hassan Awadallah, Aidan C Crook, Imed Zitouni, and Tasos Anastasakos. Understanding user satisfaction with intelligent assistants. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*, pages 121–130. ACM, 2016.

[234] Bart P Knijnenburg, Alfred Kobsa, and Hongxia Jin. Dimensionality of information disclosure behavior. *International Journal of Human-Computer Studies*, 71(12):1144–1162, 2013.

[235] Bart P Knijnenburg, Alfred Kobsa, and Hongxia Jin. Preference-based location sharing: are more privacy options really better? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2667–2676, 2013.

[236] Rafal Kocielnik, Daniel Avrahami, Jennifer Marlow, Di Lu, and Gary Hsieh. Designing for workplace reflection: a chat and voice-based conversational agent. In *Proceedings of the 2018 Designing Interactive Systems Conference*, pages 881–894, 2018.

[237] Spyros Kokolakis. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, 64:122–134, 2017.

[238] Sherrie YX Komiak and Izak Benbasat. The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS quarterly*, pages 941–960, 2006.

[239] Gerd Kortuem, Zary Segall, and Martin Bauer. Context-aware, adaptive wearable computers as remote interfaces to'intelligent'environments. In *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No. 98EX215)*, pages 58–65. IEEE, 1998.

[240] Sylvia Kowalewski, Martina Ziefle, Henrik Ziegeldorf, and Klaus Wehrle. Like us on facebook!–analyzing user preferences regarding privacy settings in germany. *Procedia Manufacturing*, 3:815–822, 2015.

[241] Robert V Kozinets. The field behind the screen: Using netnography for marketing research in online communities. *Journal of marketing research*, 39(1):61–72, 2002.

[242] Hanna Krasnova, Sarah Spiekermann, Ksenia Koroleva, and Thomas Hildebrand. Online social networks: Why we disclose. *Journal of information technology*, 25(2):109–125, 2010.

[243] Balachander Krishnamurthy and Craig E Wills. Characterizing privacy in online social networks. In *Proceedings of the first workshop on Online social networks*, pages 37–42, 2008.

[244] Balachander Krishnamurthy and Craig E Wills. On the leakage of personally identifiable information via online social networks. In *Proceedings of the 2nd ACM workshop on Online social networks*, pages 7–12, 2009.

[245] Annica Kristoffersson, Silvia Coradeschi, and Amy Loutfi. A review of mobile robotic telepresence. *Advances in Human-Computer Interaction*, 2013:3, 2013.

[246] Priya Kumar and Sarita Schoenebeck. The modern day baby book: Enacting good mothering and stewarding privacy on facebook. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 1302–1312. ACM, 2015.

[247] Ponnurangam Kumaraguru and Lorrie Faith Cranor. *Privacy indexes: a survey of Westin's studies*. Carnegie Mellon University, School of Computer Science, Institute for ..., 2005.

[248] Anastasia Kuzminykh and Jessica R Cauchard. Be mine: Contextualization ofownership research in hci. In *Proceedings of the 2020 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2020.

[249] Anastasia Kuzminykh and Jessica R Cauchard. How we own drones: On the sense of ownership in the drone design. In *iHDI@ CHI*, 2020.

[250] Anastasia Kuzminykh and Edward Lank. People searched by people: Context-based selectiveness in online search. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, pages 749–760. ACM, 2016.

[251] Anastasia Kuzminykh and Edward Lank. How much is too much? understanding the information needs of parents of young children. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(2):1–21, 2019.

[252] Anastasia Kuzminykh and Sean Rintel. Classification of functional attention in video meetings. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[253] Anastasia Kuzminykh and Sean Rintel. Low engagement as a deliberate practice of remote participants in video meetings. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–9, 2020.

[254] Anastasia Kuzminykh, Jenny Sun, Nivetha Govindaraju, Jeff Avery, and Edward Lank. Genie in the bottle: Anthropomorphized perceptions of conversational agents. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[255] Anastasia Kuzminykh, James Wallace, Richard Zanibbi, Stacey Scott, and Edward Lank. How changing communication channels affect communication patterns: Implications for the design of smart objects.

[256] Anastasia A Kuzminykh and Sergey N Enikolopov. Perception of aggressiveness of communicative information depending on a speaker's gender. *International Journal of Information and Education Technology*, 5(1):68, 2015.

[257] Cornell Social Media Lab. What makes us love alexa?, 1 2017. Retrieved February 28, 2019 from https://medium.com/social-media-stories/what-makes-us-love-alexa-dd86e61be936.

[258] Taranu I. Hartenstein H. Labitzke, S. What your friends tell others about you: Low cost linkability of social network profiles. In *Proc. 5th International ACM Workshop on Social Network Mining and Analysis*. ACM Press, 2011.

[259] Airi Lampinen, Vilma Lehtinen, Asko Lehmuskallio, and Sakari Tamminen. We're in it together: interpersonal management of disclosure in social network services. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 3217–3226, 2011.

[260] Patricia G Lange. Publicly private and privately public: Social networking on youtube. *Journal of computer-Mediated communication*, 13(1):361–380, 2007.

[261] Nancy K Lankton, D Harrison McKnight, and John F Tripp. Facebook privacy management strategies: A cluster analysis of user privacy behaviors. *Computers in Human Behavior*, 76:149–163, 2017.

[262] Liliana Laranjo, Adam G Dunn, Huong Ly Tong, Ahmet Baki Kocaballi, Jessica Chen, Rabia Bashir, Didi Surian, Blanca Gallego, Farah Magrabi, Annie YS Lau, et al. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, 25(9):1248–1258, 2018.

[263] Reed Larson and Mihaly Csikszentmihalyi. The experience sampling method. *New Directions for Methodology of Social & Behavioral Science*, 1983.

[264] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. Alexa, are you listening? privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–31, 2018.

[265] Alessandro Acquisti Laura Brandimarte and George Loewenstein. Misplaced Confidences Privacy and the Control Paradox. *Social Psychological and Personality Science*, 4(3):340–347, 2013.

[266] Kiron Lebeck, Kimberly Ruth, Tadayoshi Kohno, and Franziska Roesner. Towards security and privacy for multi-user augmented reality: Foundations with end users. In *2018 IEEE Symposium on Security and Privacy (SP)*, pages 392–408. IEEE, 2018.

[267] SeoYoung Lee and Junho Choi. Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *International Journal of Human-Computer Studies*, 103:95–105, 2017.

[268] Marvin Leibowitz. *Interpreting projective drawings: A self-psychological approach*. Routledge, 2013.

[269] Séverin Lemaignan, Julia Fink, and Pierre Dillenbourg. The dynamics of anthropomorphism in robotics. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 226–227. IEEE, 2014.

[270] Aleksei N Leont'ev. The problem of activity in psychology. *Soviet psychology*, 13(2):4–33, 1974.

[271] MR Levin. Amazon echo: What we know now. *The Huffington Post*, 2016.

[272] Toby Jia-Jun Li, Igor Labutov, Brad A Myers, Amos Azaria, Alexander I Rudnicky, and Tom M Mitchell. An end user development approach for failure handling in goal-oriented conversational agents. *Studies in Conversational UX Design*, 2018.

[273] Joseph CR Licklider. Man-computer symbiosis. *IRE transactions on human factors in electronics*, (1):4–11, 1960.

[274] Ruoyun Lin and Sonja Utz. Self-disclosure on sns: Do disclosure intimacy and narrativity influence interpersonal closeness and social attraction? *Computers in Human Behavior*, 70:426–436, 2017.

[275] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.

[276] Richard Harper Lindley, Siân E. and Abigail Sellen. Designing a technological playground: a field study of the emergence of play in household messaging. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2010.

[277] Siân E Lindley, Catherine C Marshall, Richard Banks, Abigail Sellen, and Tim Regan. Rethinking the web as a personal archive. In *Proceedings of the 22nd international conference on World Wide Web*, pages 749–760, 2013.

[278] Eden Litt. Knock, knock. who's there? the imagined audience. *Journal of broadcasting & electronic media*, 56(3):330–345, 2012.

[279] Eden Litt and Eszter Hargittai. Smile, snap, and share? a nuanced approach to privacy and online photo-sharing. *Poetics*, 42:1–21, 2014.

[280] Yabing Liu, Krishna P Gummadi, Balachander Krishnamurthy, and Alan Mislove. Analyzing facebook privacy settings: user expectations vs. reality. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, pages 61–70, 2011.

[281] Pierre Livet and Pauline Ridel. The intentions of communication. *Réseaux. Communication-Technologie-Société*, 2(1):151–175, 1994.

[282] Mascheroni G. Dreier M. Chaudron S. Livingstone, S. and K. Lagae. *How parents of young children manage digital devices at home: The role of income, education and parental style*. London: EU Kids Online, LSE, 2015.

[283] BabyCenter LLC. Baby center blogs, 2017.

[284] BabyCenter LLC. Baby center community, 2017.

[285] Irene Lopatovska, Katrina Rink, Ian Knight, Kieran Raines, Kevin Cosenza, Harriet Williams, Perachya Sorsche, David Hirsch, Qi Li, and Adrianna Martinez. Talk to me: Exploring user interactions with the amazon alexa. *Journal of Librarianship and Information Science*, page 0961000618759414, 2018.

[286] Irene Lopatovska and Harriet Williams. Personification of the amazon alexa: Bff or a mindless companion. In *Proceedings of the 2018 Conference on Human Information Interaction&Retrieval*, pages 265–268. ACM, 2018.

[287] Lori Kido Lopez. The radical act of'mommy blogging': redefining motherhood through the blogosphere. *New media & society*, 11(5):729–747, 2009.

[288] Paul Luff, Christian Heath, Hideaki Kuzuoka, Jon Hindmarsh, Keiichi Yamazaki, and Shinya Oyama. Fractured ecologies: creating environments for collaboration. *Human-Computer Interaction*, 18(1):51–84, 2003.

[289] Ewa Luger and Abigail Sellen. Like having a really bad pa: the gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5286–5297. ACM, 2016.

[290] Ziyu Lv, Feng Xia, Guowei Wu, Lin Yao, and Zhikui Chen. icare: a mobile health monitoring system for the elderly. In *2010 IEEE/ACM Int'l Conference on Green Computing and Communications & Int'l Conference on Cyber, Physical and Social Computing*, pages 699–705. IEEE, 2010.

[291] Karl F MacDorman and Debaleena Chattopadhyay. Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition*, 146:190–205, 2016.

[292] Karl F MacDorman and Hiroshi Ishiguro. The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies*, 7(3):297–337, 2006.

[293] Haley MacLeod, Cynthia L Bennett, Meredith Ringel Morris, and Edward Cutrell. Understanding blind people's experiences with computer-generated captions of social media images. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 5988–5999, 2017.

[294] Poornima Madhavan and Douglas A Wiegmann. Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science*, 8(4):277–301, 2007.

[295] Jens-Erik Mai. The quality and qualities of information. *Journal of the American society for information science and technology*, 64(4):675–688, 2013.

[296] Jens-Erik Mai. Personal information as communicative acts. *Ethics and Information Technology*, 18(1):51–57, 2016.

[297] Lydia Manikonda, Aditya Deotale, and Subbarao Kambhampati. What's up with privacy?: User preferences and privacy concerns in intelligent personal assistants. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, pages 229–235. ACM, 2018.

[298] Stephen T Margulis. Three theories of privacy: An overview. In *Privacy online*, pages 9–17. Springer, 2011.

[299] et al. Markopoulos, Panos. Keeping in touch with the family: home and away with the astra awareness system. In *CHI'04 extended abstracts on Human factors in computing systems*. ACM, 2004.

[300] Catherine C Marshall and Siân E Lindley. Searching for myself: motivations and strategies for self-search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3675–3684, 2014.

[301] Catherine C Marshall and Frank M Shipman. Exploring the ownership and persistent value of facebook content. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 712–723, 2015.

[302] Kirsten Martin and Katie Shilton. Putting mobile application privacy in context: An empirical study of user privacy expectations for mobile devices. *The Information Society*, 32(3):200–216, 2016.

[303] Gary Marx and Valerie Steeves. From the beginning: Children as subjects and agents of surveillance. *Surveillance & Society*, 7(3/4):192–230, 2010.

[304] Olivier Morin Mascaro, Olivier and Dan Sperber. Optimistic expectations about communication explain children's difficulties in hiding, lying, and mistrusting liars. *Journal of child language*, 44(5):1041–1064, 2017.

[305] Dominic W Massaro, Michael M Cohen, Sharon Daniel, and Ronald A Cole. Developing and evaluating conversational agents. In *Human performance and ergonomics*, pages 173–194. Elsevier, 1999.

[306] Antony Mayfield. *Me and my web shadow: How to manage your reputation online*. A&C Black, 2010.

[307] Karen McCullagh. Blogging: self presentation and privacy. *Information & communications technology law*, 17(1):3–23, 2008.

[308] Brandon T McDaniel, Sarah M Coyne, and Erin K Holmes. New mothers and media use: Associations between blogging, social networking, and maternal well-being. *Maternal and child health journal*, 16(7):1509–1517, 2012.

[309] Albert Mehrabian et al. *Silent messages*, volume 8. Wadsworth Belmont, CA, 1971.

[310] Adam S Miner, Arnold Milstein, Stephen Schueller, Roshini Hegde, Christina Mangurian, and Eleni Linos. Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA internal medicine*, 176(5):619–625, 2016.

[311] AJung Moon, Daniel M Troniak, Brian Gleeson, Matthew KXJ Pan, Minhua Zheng, Benjamin A Blumer, Karon MacLean, and Elizabeth A Croft. Meet me where i'm gazing: how shared attention gaze affects human-robot handover timing. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pages 334–341. ACM, 2014.

[312] Ana Hilda Morales-Aranda and Oscar Mayora-Ibarra. A context sensitive public display for adaptive multi-user information visualization. In *Third International Conference on Autonomic and Autonomous Systems (ICAS'07)*, pages 63–63. IEEE, 2007.

[313] Philip R Cohen Jerry L Morgan and Martha E Pollack. *Intentions in communication*. MIT press, 1990.

[314] Anthony Morton and M Angela Sasse. Desperately seeking assurances: Segmenting users by their information-seeking preferences. In *2014 Twelfth Annual International Conference on Privacy, Security and Trust*, pages 102–111. IEEE, 2014.

[315] Carol Moser, Tianying Chen, and Sarita Y Schoenebeck. Parents' and children's preferences about parents sharing about children on social media. In *Proceedings of*

the 2017 CHI Conference on Human Factors in Computing Systems, pages 5221–5225. ACM, 2017.

[316] Peter Mundy and Lisa Newell. Attention, joint attention, and social cognition. *Current directions in psychological science*, 16(5):269–274, 2007.

[317] Miguel A Muñoz, Victor M Gonzalez, Marcela Rodríguez, and Jesus Favela. Supporting context-aware collaboration in a hospital: an ethnographic informed design. In *International Conference on Collaboration and Technology*, pages 330–344. Springer, 2003.

[318] Kevin Murnane. How older and younger millennials differ in their approach to online privacy and security. https://www.forbes.com/sites/kevinmurnane/2016/04/13/how-older-and-younger-millennials-differ-in-their-approach-to-online-privacy-and-security/#6952d0699aa3, 2016.

[319] Elizabeth D. Mynatt, Jim Rowan, Sarah Craighill, and Annie Jacobs. Digital family portraits: Supporting peace of mind for extended family members. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pages 333–340, New York, NY, USA, 2001. ACM.

[320] Guillaume Nadon, Marcus Feilberg, Mathias Johansen, and Irina Shklovski. In the user we trust: Unrealistic expectations of facebook's privacy mechanisms. In *Proceedings of the 9th International Conference on Social Media and Society*, pages 138–149, 2018.

[321] Kris Nagel, Cory D Kidd, Thomas O'Connell, Anind Dey, and Gregory D Abowd. The family intercom: Developing a context-aware audio communication system. In *International Conference on Ubiquitous Computing*, pages 176–183. Springer, 2001.

[322] Kristine S Nagel, James M Hudson, and Gregory D Abowd. Predictors of availability in home life context-mediated communication. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 497–506, 2004.

[323] Clifford Nass and Kwan Min Lee. Does computer-synthesized speech manifest personality? experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of experimental psychology: applied*, 7(3):171, 2001.

[324] Clifford Nass and Youngme Moon. Machines and mindlessness: Social responses to computers. 56(1), 2000.

[325] Clifford Nass, Youngme Moon, and Nancy Green. Are machines gender neutral? gender-stereotypic responses to computers with voices. *Journal of applied social psychology*, 27(10):864–876, 1997.

[326] Clifford Nass, Jonathan Steuer, and Ellen R Tauber. Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 72–78. ACM, 1994.

[327] Thomas Neumayr, Hans-Christian Jetter, Mirjam Augstein, Judith Friedl, and Thomas Luger. Domino: A descriptive framework for hybrid collaboration and coupling styles in partially distributed teams. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–24, 2018.

[328] Carman Neustaedter and A. J. Bernheim Brush. Linc-ing the family: the participatory design of an inkable family calendar. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006.

[329] Kathryn Elliot Neustaedter, Carman and Saul Greenberg. Interpersonal awareness in the domestic realm. In *Proceedings of the 18th Australia conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments*. ACM, 2006.

[330] Paul Newland, James A Powell, and Chris Creed. Understanding architectural designers' selective information handling. *Design Studies*, 8(1):2–16, 1987.

[331] Guardian News and Media Limited. The guardian, 2017.

[332] MIT News. Mit web site gathers and displays information on government, 2003. Accessed: 2020-06-02.

[333] Erik G Nilsson, Jacqueline Floch, Svein Hallsteinsen, and Erlend Stav. Model-based user interface adaptation. *Computers & Graphics*, 30(5):692–701, 2006.

[334] Helen Nissenbaum. A contextual approach to privacy online. *Daedalus*, 140(4):32–48, 2011.

[335] Dongfang Niu, Jacques Terken, and Berry Eggen. Anthropomorphizing information to enhance trust in autonomous vehicles. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 28(6):352–359, 2018.

[336] Oded Nov, Mor Naaman, and Chen Ye. Motivational, structural and tenure factors that impact online community photo sharing. In *Third International AAAI Conference on Weblogs and Social Media*, 2009.

[337] Brid O'Conaill, Steve Whittaker, and Sylvia Wilbur. Conversations over video conferences: An evaluation of the spoken aspects of video-mediated communication. *Human-computer interaction*, 8(4):389–428, 1993.

[338] Jean Oh, Felipe Meneguzzi, and Katia Sycara. Probabilistic plan recognition for intelligent information agents. *Proc. ICAART*, pages 281–287, 2011.

[339] Aude Oliva and Antonio Torralba. The role of context in object recognition. *Trends in cognitive sciences*, 11(12):520–527, 2007.

[340] Judith S Olson, Jonathan Grudin, and Eric Horvitz. A study of preferences for sharing and privacy. In *CHI'05 extended abstracts on Human factors in computing systems*, pages 1985–1988, 2005.

[341] J O'Neill. The disturbing facebook trend of stolen kids photos. https://www.yahoo.com/news/mom-my-son-was-digitally-kidnapped-what-112545291567.html, 2015.

[342] Sheryl Ong and Aaron Suplizio. Unpacking the breakout success of the amazon echo. *Retrieved December*, 12:2018, 2016.

[343] D. K. Orthner and J. A. Mancini. Benefits of leisure experiences for family bonding. In *Benefits of leisure*. State College, PA: Vantage, 1991.

[344] Babajide Osatuyi. Is lurking an anxiety-masking strategy on social media sites? the effects of lurking and computer anxiety on explaining information privacy concern on social media platforms. *Computers in Human Behavior*, 49:324–332, 2015.

[345] Mai Otsuki, Taiki Kawano, Keita Maruyama, Hideaki Kuzuoka, and Yusuke Suzuki. Thirdeye: Simple add-on display to represent remote participant's gaze direction in video communication. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 5307–5312. ACM, 2017.

[346] Mai Otsuki, Keita Maruyama, Hideaki Kuzuoka, and Yusuke Suzuki. Effects of enhanced gaze presentation on gaze leading in remote collaborative physical tasks. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 368. ACM, 2018.

[347] L. Palen and A. Hughes. When home base is not a place: Parents use of mobile telephones. *Personal and Ubiquitous Computing*, 11(5):339–348, 2007.

[348] Ye Pan and Anthony Steed. Preserving gaze direction in teleconferencing using a camera array and a spherical display. In *2012 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*, pages 1–4. IEEE, 2012.

[349] Zizi Papacharissi. The virtual geographies of social networks: a comparative analysis of facebook, linkedin and asmallworld. *New media & society*, 11(1-2):199–220, 2009.

[350] Raja Parasuraman and Dietrich H Manzey. Complacency and bias in human use of automation: An attentional integration. *Human factors*, 52(3):381–410, 2010.

[351] J. Pascoe. Adding generic contextual capabilities to wearable computers. In *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No.98EX215)*, pages 92–99, Oct 1998.

[352] Sarah Perez. Alexa skills top 80,000 after a big alexa-powered holiday season, 2019.

[353] Dr Marta Perez Garcia, Sarita Saffon Lopez, and Hector Donis. Everybody is talking about virtual assistants, but how are people really using them? In *Proceedings of the 32nd International BCS Human Computer Interaction Conference 32*, pages 1–5, 2018.

[354] John Durham Peters. Information: Notes toward a critical history. *Journal of Communication Inquiry*, 12(2):9–23, 1988.

[355] Sandra Petronio. *Boundaries of privacy: Dialectics of disclosure*. Suny Press, 2002.

[356] Sandra Petronio. Brief status report on communication privacy management theory. *Journal of Family Communication*, 13(1):6–14, 2013.

[357] Rosalind W Picard. Affective computing. *The MIT Press, Cambridge (MA)*, 167:170, 1997.

[358] Krzysztof Pietroszek, Anastasia Kuzminykh, James R Wallace, and Edward Lank. Smartcasting: a discount 3d interaction technique for public displays. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design*, pages 119–128, 2014.

[359] David Pinelle and Carl Gutwin. A groupware design framework for loosely coupled workgroups. In *ECSCW 2005*, pages 65–82. Springer, 2005.

[360] Jean A. Pratt, Karina Hauser, Zsolt Ugray, and Olga Patterson. Looking at human-computer interface design: Effects of ethnicity in computer agents. *Interact. Comput.*, 19(4):512–523, July 2007.

[361] Paul Prekop and Mark Burnett. Activities, context and ubiquitous computing. *Computer communications*, 26(11):1168–1176, 2003.

[362] Aina Puce and Bennett I Bertenthal. *The Many Faces of Social Attention: Behavioral and Neural Measures*. Springer, 2015.

[363] Amanda Purington, Jessie G. Taft, Shruti Sannon, Natalya N. Bazarova, and Samuel Hardman Taylor. "alexa is my new bff": Social roles, user satisfaction, and personification of the amazon echo. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, pages 2853–2859, New York, NY, USA, 2017. ACM.

[364] Amanda Purington, Jessie G Taft, Shruti Sannon, Natalya N Bazarova, and Samuel Hardman Taylor. Alexa is my new bff: social roles, user satisfaction, and personification of the amazon echo. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 2853–2859. ACM, 2017.

[365] Lingyun Qiu and Izak Benbasat. Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems. *Journal of Management Information Systems*, 25(4):145–182, 2009.

[366] Kelly Quinn. Why we share: A uses and gratifications approach to privacy regulation in social media use. *Journal of Broadcasting & Electronic Media*, 60(1):61–86, 2016.

[367] Emilee Rader and Janine Slaker. The importance of visibility for folk theories of sensor data. In *Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017)*, pages 257–270, 2017.

[368] Kistin C. J. Zuckerman B. Nitzberg K. Gross J. Kaplan-Sanoff M. et al. Radesky, J. S. Patterns of mobile device use by caregivers and children during meals in fast food restaurants. *Pediatrics*, 133(4):843–850, 2014.

[369] Irene Rae, Gina Venolia, John C Tang, and David Molnar. A framework for understanding and designing telepresence. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, pages 1552–1566. ACM, 2015.

[370] Pauline Rafferty and Rob Hidderley. *Indexing multimedia and creative works: the problems of meaning and interpretation*. Routledge, 2017.

[371] Hayes Raffle, Glenda Revelle, Koichi Mori, Rafael Ballagas, Kyle Buza, Hiroshi Horii, Joseph Kaye, Kristin Cook, Natalie Freed, Janet Go, and Mirjana Spasojevic. Hello, is grandma there? let's read! storyvisit: Family video chat and connected e-books. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1195–1204, New York, NY, USA, 2011. ACM.

[372] Vignesh Ramanathan, Jonathan Huang, Sami Abu-El-Haija, Alexander Gorban, Kevin Murphy, and Li Fei-Fei. Detecting events and key actors in multi-person videos. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3043–3053, 2016.

[373] Vignesh Ramanathan, Bangpeng Yao, and Li Fei-Fei. Social role discovery in human events. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2475–2482, 2013.

[374] Anand Ranganathan, Roy H Campbell, Arathi Ravi, and Anupama Mahajan. Conchat: A context-aware chat program. *IEEE Pervasive computing*, 1(3):51–57, 2002.

[375] Shanaka Ransiri and Suranga Nanayakkara. Watchme: Wrist-worn interface that makes remote monitoring seamless. In *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '12, pages 243–244, New York, NY, USA, 2012. ACM.

[376] Judith Rauhofer. 'look to yourselves, that we lose not those things which we have wrought.'what do the proposed changes to the purpose limitation principle mean for public bodies' rights to access third-party data? *International Review of Law, Computers & Technology*, 28(2):144–158, 2014.

[377] Reza Rawassizadeh. Towards sharing life-log information with society. *Behaviour & Information Technology*, 31(11):1057–1067, 2012.

[378] Erika Rawes. The funniest things to ask alexa, 5 2020. Retrieved May 25, 2020 from https://www.digitaltrends.com/home/funny-things-to-ask-alexa/.

[379] Maryam N Razavi and Lee Iverson. A grounded theory of information sharing behavior in a personal learning space. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 459–468, 2006.

[380] reddit inc. Reddit fitbit, 2017.

[381] reddit inc. Reddit parenting, 2017.

[382] Stuart Reeves. Envisioning ubiquitous computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1573–1582. ACM, 2012.

[383] A. Reis, D. Paulino, H. Paredes, I. Barroso, M. J. Monteiro, V. Rodrigues, and J. Barroso. Using intelligent personal assistants to assist the elderlies an evaluation of amazon alexa, google assistant, microsoft cortana, and apple siri. In *2018 2nd International Conference on Technology and Innovation in Sports, Health and WellFbeing (TISHW)*, pages 1–5, June 2018.

[384] Reputation.com. Accessed: 2020-06-02.

[385] Bolivar Reyes-Jaquez and Catharine H. Echols. Playing by the rules: Self-interest information influences children's trust and trustworthiness in the absence of feedback. *Cognition*, 134:140–154, 2015.

[386] Bernardo Reynolds, Jayant Venkatanathan, Jorge Gonçalves, and Vassilis Kostakos. Sharing ephemeral information in online social networks: privacy perceptions and behaviours. In *IFIP Conference on Human-Computer Interaction*, pages 204–215. Springer, 2011.

[387] Daniel C Richardson, Rick Dale, and John M Tomlinson. Conversation, gaze coordination, and beliefs about visual context. *Cognitive Science*, 33(8):1468–1482, 2009.

[388] Laurel D Riek, Tal-Chen Rabinowitch, Bhismadev Chakrabarti, and Peter Robinson. How anthropomorphism affects empathy toward robots. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, pages 245–246. ACM, 2009.

[389] Sean Rintel. Mobile sharing and companion experiences for microsoft teams meetings, 3 2018. Retrieved September 07, 2019 from https://www.microsoft.com/en-us/garage/wall-of-fame/companionexperiences/.

[390] Evan F Risko, Daniel C Richardson, and Alan Kingstone. Breaking the fourth wall of cognitive science: Real-world social attention and the dual function of gaze. *Current Directions in Psychological Science*, 25(1):70–74, 2016.

[391] Ben Robins, Paul Dickerson, Penny Stribling, and Kerstin Dautenhahn. Robot-mediated joint attention in children with autism: A case study in robot-human interaction. *Interaction studies*, 5(2):161–198, 2004.

[392] John Robinson, Sarah Burch, Sonia Talwar, Meg O'Shea, and Mike Walsh. Envisioning sustainability: Recent progress in the use of participatory backcasting approaches for sustainability research. *Technological Forecasting and Social Change*, 78(5):756–768, 2011.

[393] Finlo Rohrer. When all video all, 2009. Accessed: 2020-06-02.

[394] et al. Romero, Natalia. Connecting the family with awareness systems. *Personal and Ubiquitous Computing*, 11(4):299–312, 2007.

[395] Ariel Rosenfeld, Noa Agmon, Oleg Maksimov, and Sarit Kraus. Intelligent agent supporting human–multi-robot team collaboration. *Artificial Intelligence*, 252:211–231, 2017.

[396] Jim Rowan and Elizabeth D. Mynatt. Digital family portrait field trial: Support for aging in place. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 521–530, New York, NY, USA, 2005. ACM.

[397] Ariel Rubinstein. "economics and psychology"? the case of hyperbolic discounting. *International Economic Review*, 44(4):1207–1216, 2003.

[398] Daniel M Russell, Clemens Drews, and Alison Sue. Social aspects of using large public interactive displays for collaboration. In *International Conference on Ubiquitous Computing*, pages 229–236. Springer, 2002.

[399] Kathy Ryall, Clifton Forlines, Chia Shen, and Meredith Ringel Morris. Exploring the effects of group size and table size on interactions with tabletop shared-display groupware. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 284–293, 2004.

[400] Sara Saeedi, Adel Moussa, and Naser El-Sheimy. Context-aware personal navigation using embedded sensor fusion in smartphones. *Sensors*, 14(4):5742–5767, 2014.

[401] Safewise. The 15 best gps kids trackers for parents with young kids, 2017.

[402] Kambiz Saffarizadeh, Maheshwar Boodraj, and Tawfiq M Alashoor. Conversational assistants: Investigating privacy concerns, trust, and self-disclosure. 2017.

[403] Günther Sagl, Bernd Resch, Andreas Petutschnig, Kalliopi Kyriakou, Michael Liedlgruber, and Frank H Wilhelm. Wearables and the quantified self: Systematic benchmarking of physiological sensors. *Sensors*, 19(20):4448, 2019.

[404] Hanan Salam and Mohamed Chetouani. A multi-level context-based modeling of engagement in human-robot interaction. In *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, volume 3, pages 1–6. IEEE, 2015.

[405] Brenda Salley and John Colombo. Conceptualizing social attention in developmental research. *Social Development*, 25(4):687–703, 2016.

[406] Holger Salzmann, Mathias Moehring, and Bernd Froehlich. Virtual vs. real-world pointing in two-user scenarios. In *2009 IEEE Virtual Reality Conference*, pages 127–130. IEEE, 2009.

[407] Samiha Samrose, Ru Zhao, Jeffery White, Vivian Li, Luis Nova, Yichen Lu, Mohammad Rafayet Ali, and Mohammed Ehsan Hoque. Coco: Collaboration coach for understanding team dynamics during video conferencing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(4):160, 2018.

[408] Manuel J Sánchez-Franco and José L Roldán. The influence of familiarity, trust and norms of reciprocity on an experienced sense of community: an empirical analysis based on social online services. *Behaviour & Information Technology*, 34(4):392–412, 2015.

[409] B. Schilit, N. Adams, and R. Want. Context-aware computing applications. In *Proceedings of the 1994 First Workshop on Mobile Computing Systems and Applications*, WMCSA '94, pages 85–90, Washington, DC, USA, 1994. IEEE Computer Society.

[410] Bill Schilit, Norman Adams, and Roy Want. Context-aware computing applications. In *1994 First Workshop on Mobile Computing Systems and Applications*, pages 85–90. IEEE, 1994.

[411] Bill N Schilit, Norman Adams, Rich Gold, Michael M Tso, and Roy Want. The parctab mobile computing system. In *Proceedings of IEEE 4th Workshop on Workstation Operating Systems. WWOS-III*, pages 34–39. IEEE, 1993.

[412] Bill N Schilit, David M Hilbert, and Jonathan Trevor. Context-aware communication. *IEEE Wireless Communications*, 9(5):46–54, 2002.

[413] Albrecht Schmidt, Michael Beigl, and Hans-W Gellersen. There is more to context than location. *Computers & Graphics*, 23(6):893–901, 1999.

[414] Simeon Schudy and Verena Utikal. 'you must not know about me'—on the willingness to share personal data. *Journal of Economic Behavior & Organization*, 141:1–13, 2017.

[415] Alex Sciuto, Arnita Saini, Jodi Forlizzi, and Jason I. Hong. "hey alexa, what's up?": A mixed-methods studies of in-home conversational agent usage. In *Proceedings of the 2018 Designing Interactive Systems Conference*, DIS '18, pages 857–868, New York, NY, USA, 2018. ACM.

[416] Alex Sciuto, Arnita Saini, Jodi Forlizzi, and Jason I Hong. Hey alexa, what's up?: A mixed-methods studies of in-home conversational agent usage. In *Proceedings of the 2018 on Designing Interactive Systems Conference 2018*, pages 857–868. ACM, 2018.

[417] Stacey D Scott, M Sheelagh T Carpendale, and Kori M Inkpen. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 294–303, 2004.

[418] Stacey D Scott and Sheelagh Carpendale. Investigating tabletop territoriality in digital tabletop workspaces. 2006.

[419] John R Searle and John Rogers Searle. *Speech acts: An essay in the philosophy of language*, volume 626. Cambridge university press, 1969.

[420] Ted Selker, Andrea Lockerd, and Jorge Martinez. Eye-r, a glasses-mounted eye motion detection interface. In *CHI'01 extended abstracts on Human factors in computing systems*, pages 179–180. ACM, 2001.

[421] Abigail J Sellen. Remote conversations: The effects of mediating talk with technology. *Human-computer interaction*, 10(4):401–444, 1995.

[422] Common Sense. Common sense media, 2017.

[423] Kalpana Shankar, L Jean Camp, Kay Connelly, and Lesa Huber. Aging, privacy, and home-based computing: Developing a design framework. *IEEE Pervasive Computing*, 11(4):46–54, 2012.

[424] Claude E Shannon and Warren Weaver. The mathematical theory of information. *Urbana: University of Illinois Press*, 97, 1949.

[425] Claude Elwood Shannon. Communication in the presence of noise. *Proceedings of the IEEE*, 86(2):447–457, 1998.

[426] Kim Bartel Sheehan. Toward a typology of internet users and online privacy concerns. *The Information Society*, 18(1):21–32, 2002.

[427] Samira Sheikhi, Dinesh Babu Jayagopi, Vasil Khalidov, and Jean-Marc Odobez. Context aware addressee estimation for human robot interaction. In *Proceedings of the 6th workshop on Eye gaze in intelligent human machine interaction: gaze in multimodal interaction*, pages 1–6. ACM, 2013.

[428] John Short, Ederyn Williams, and Bruce Christie. The social psychology of telecommunications. 1976.

[429] Garriy Shteynberg. A social host in the machine? the case of group attention. *Journal of Applied Research in Memory and Cognition*, 3(4):307–311, 2014.

[430] Garriy Shteynberg. Shared attention. *Perspectives on Psychological Science*, 10(5):579–590, 2015.

[431] Simple Simon and Company. 25 ways to ask your kids "so how was school today?" without asking them "so how was school today?", 2017.

[432] Manya Sleeper, Rebecca Balebako, Sauvik Das, Amber Lynn McConahy, Jason Wiese, and Lorrie Faith Cranor. The post that wasn't: exploring self-censorship on facebook. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 793–802, 2013.

[433] Joanne R Smith, Deborah J Terry, and Michael A Hogg. Who will see me? the impact of type of audience on willingness to display group-mediated attitude–intention consistency. *Journal of Applied Social Psychology*, 36(5):1173–1197, 2006.

[434] Scott S Snibbe and Hayes S Raffle. Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1447–1456, 2009.

[435] Beate Sodian. The development of deception in young children. *British Journal of Developmental Psychology*, 9(1):173–188, 1991.

[436] Daniel J Solove. Understanding privacy. 2008.

[437] Makram Soui, Soumaya Diab, Ali Ouni, Aroua Essayeh, and Mourad Abed. An ontology-based approach for user interface adaptation. In *Advances in Intelligent Systems and Computing*, pages 199–215. Springer, 2017.

[438] Abigale Stangl, Meredith Ringel Morris, and Danna Gurari. " person, shoes, tree. is the person naked?" what people with vision impairments want in image descriptions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[439] statista.com. Number of social media users worldwide from 2010 to 2021. https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/, 2017.

[440] Häkan Stattin and Margaret Kerr. Parental monitoring: A reinterpretation. *Child Development*, 71(4):1072–1085, 2000.

[441] Maria Staudte and Matthew W Crocker. Visual attention in spoken human-robot interaction. In *Human-Robot Interaction (HRI), 2009 4th ACM/IEEE International Conference on*, pages 77–84. IEEE, 2009.

[442] Lee Stearns and Anja Thieme. Automated person detection in dynamic scenes to assist people with vision impairments: An initial investigation. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 391–394, 2018.

[443] Michael A Stefanone and Derek Lackaff. Reality television as a model for online behavior: Blogging, photo, and video sharing. *Journal of computer-mediated communication*, 14(4):964–987, 2009.

[444] Stacey B Steinberg. Sharenting: Children's privacy in the age of social media. *Emory LJ*, 66:839, 2016.

[445] Keli Ryan Steuber and Rachel M McLaren. Privacy recalibration in personal relationships: Rule usage before and after an incident of privacy turbulence. *Communication Quarterly*, 63(3):345–364, 2015.

[446] Rainer Stiefelhagen, Jie Yang, and Alex Waibel. Estimating focus of attention based on gaze and sound. In *Proceedings of the 2001 workshop on Perceptive user interfaces*, pages 1–9. ACM, 2001.

[447] Thomas Strang and Claudia Linnhoff-Popien. A context modeling survey. In *Workshop Proceedings*, 2004.

[448] Katherine Strater and Heather Richter Lipford. Strategies and struggles with privacy in an online social networking community. In *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction-Volume 1*, pages 111–119. British Computer Society, 2008.

[449] Yolande Strengers. Envisioning the smart home: reimagining a smart energy future. *Digital materialities: Design and anthropology*, 61:61–76, 2016.

[450] Allyson Haynes Stuart. Google search results: buried if not forgotten. *NCJL & Tech.*, 15:463, 2013.

[451] Fred Stutzman and Jacob Kramer-Duffield. Friends only: examining a privacy-enhancing behavior in facebook. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1553–1562, 2010.

[452] Kaveri Subrahmanyam, Patricia M Greenfield, and Brendesha Tynes. Constructing sexuality and identity in an online teen chat room. *Journal of applied developmental psychology*, 25(6):651–666, 2004.

[453] Gita Sukthankar, Christopher Geib, Hung Hai Bui, David Pynadath, and Robert P Goldman. *Plan, activity, and intent recognition: Theory and practice*. Newnes, 2014.

[454] Yasuyuki Sumi, Tameyuki Etani, Sidney Fels, Nicolas Simonet, Kaoru Kobayashi, and Kenji Mase. C-map: Building a context-aware mobile assistant for exhibition tours. In *Community computing and support systems*, pages 137–154. Springer, 1998.

[455] Qianru Sun, Bernt Schiele, and Mario Fritz. A domain based approach to social relation recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3481–3490, 2017.

[456] S Shyam Sundar, Hyunjin Kang, Mu Wu, Eun Go, and Bo Zhang. Unlocking the privacy paradox: do cognitive heuristics hold the key? In *CHI'13 extended abstracts on human factors in computing systems*, pages 811–816. ACM, 2013.

[457] Stefano Taddei and Bastianina Contena. Privacy, trust and control: Which relationships with online self-disclosure? *Computers in Human Behavior*, 29(3):821–826, 2013.

[458] Monika Taddicken. The 'privacy paradox' in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication*, 19(2):248–273, 2014.

[459] John C Tang and Monica Rua. Montage: Providing teleproximity for distributed groups. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 37–43, 1994.

[460] Sarah Tausch, Stephanie Ta, and Heinrich Hussmann. A comparison of cooperative and competitive visualizations for co-located collaboration. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 5034–5039, 2016.

[461] AJ Bernheim Brush Tee, Kimberly and Kori M. Inkpen. Exploring communication and sharing between extended families. *International Journal of Human-Computer Studies*, 67(2):128–138, 2009.

[462] Kimberly Tee, Saul Greenberg, and Carl Gutwin. Providing artifact awareness to a distributed group through screen sharing. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 99–108, 2006.

[463] Stergios Tegos, Stavros Demetriadis, Pantelis M Papadopoulos, and Armin Weinberger. Conversational agents for academically productive talk: A comparison of directed and undirected agent interventions. *International Journal of Computer-Supported Collaborative Learning*, 11(4):417–440, 2016.

[464] Punnarumol Temdee and Ramjee Prasad. *Context-aware communication and computing: Applications for smart environment*. Springer, 2018.

[465] David Tennenhouse. Proactive computing. *Communications of the ACM*, 43(5):43–50, 2000.

[466] Loren G Terveen. Overview of human-computer collaboration. *Knowledge-Based Systems*, 8(2-3):67–81, 1995.

[467] James C Thompson, J Gregory Trafton, and Patrick McKnight. The perception of humanness from the movements of synthetic agents. *Perception*, 40(6):695–704, 2011.

[468] Antonio Torralba, Kevin P Murphy, William T Freeman, Mark A Rubin, et al. Context-based vision system for place and object recognition. In *ICCV*, volume 3, pages 273–280, 2003.

[469] Francois Toutain, Ahmed Bouabdallah, Radim Zemek, and Claude Daloz. Interpersonal context-aware communication services. *IEEE Communications Magazine*, 49(1):68–74, 2011.

[470] Chih-Hsiung Tu. The relationship between social presence and online privacy. *The internet and higher education*, 5(4):293–318, 2002.

[471] Joseph Turow and Lilach Nir. The internet and the family: The view from parents, the view from kids. 2000.

[472] Moshe Unger, Ariel Bar, Bracha Shapira, and Lior Rokach. Towards latent context-aware recommendation systems. *Knowledge-Based Systems*, 104:165–178, 2016.

[473] Kirsikka Vaajakallio and Tuuli Mattelmäki. Collaborative design exploration: envisioning future practices with make tools. In *Proceedings of the 2007 conference on Designing pleasurable products and interfaces*, pages 223–238. ACM, 2007.

[474] Julia van Heek, Katrin Aming, and Martina Ziefle. "how fear of crime affects needs for privacy & safety": Acceptance of surveillance technologies in smart cities. In *2016 5th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS)*, pages 1–12. IEEE, 2016.

[475] Julia van Heek, Anne Kathrin Schaar, Bianka Trevisan, Patrycja Bosowski, and Martina Ziefle. User requirements for wearable smart textiles: does the usage context matter (medical vs. sports)? In *Proceedings of the 8th international conference on pervasive computing technologies for healthcare*, pages 205–209, 2014.

[476] Nancy Van House, Marc Davis, Yuri Takhteyev, Nathan Good, Anita Wilhelm, and Megan Finn. From "what?" to "why?": the social uses of personal photos. In *Proc. of CSCW 2004*, 2004.

[477] Marie E. Vasek. Lying as a skill: The development of deception in children. *Deception: Perspectives on human and nonhuman deceit*, pages 271–292, 1986.

[478] Alladi Venkatesh. Computers and other interactive technologies for the home. *Commun. ACM*, 39(12):47–54, December 1996.

[479] Roel Vertegaal. The gaze groupware system: mediating joint attention in multiparty communication and collaboration. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 294–301. ACM, 1999.

[480] Roel Vertegaal, Jeffrey S Shell, Daniel Chen, and Aadil Mamuji. Designing for augmented attention: Towards a framework for attentive user interfaces. *Computers in Human Behavior*, 22(4):771–789, 2006.

[481] Jessica Vitak. The impact of context collapse and privacy on social network site disclosures. *Journal of broadcasting & electronic media*, 56(4):451–470, 2012.

[482] voicebot.ai. Voice assistant consumer adoption report, 2018.

[483] voicebot.ai. Smart speaker consumer adoption report, 2019.

[484] Amy Voida and Elizabeth D Mynatt. Six themes of the communicative appropriation of photographic images. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 171–180. ACM, 2005.

[485] Sarah Theres Völkel, Daniel Buschek, Jelena Pranjic, and Heinrich Hussmann. Understanding emoji interpretation through user personality and message context. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 1–12, 2019.

[486] Alexandra Vtyurina and Adam Fourney. Exploring the role of conversational cues in guided task support with virtual assistants. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–7, 2018.

[487] Amina Wagner, Lisa Alina Gasche, et al. Sharenting: Making decisions about other's privacy on social networking sites. Technical report, Darmstadt Technical University, Department of Business Administration ..., 2018.

[488] James R Wallace, Nancy Iskander, and Edward Lank. Creating your bubble: Personal space on and around large public displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2087–2092, 2016.

[489] Roy Want, Andy Hopper, Veronica Falcao, and Jonathan Gibbons. The active badge location system. *ACM Transactions on Information Systems (TOIS)*, 10(1):91–102, 1992.

[490] Adam Waytz, John Cacioppo, and Nicholas Epley. Who sees human? the stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, 5(3):219–232, 2010.

[491] Adam Waytz, Nicholas Epley, and John T Cacioppo. Social cognition unbound: Insights into anthropomorphism and dehumanization. *Current Directions in Psychological Science*, 19(1):58–62, 2010.

[492] Adam Waytz, Joy Heafner, and Nicholas Epley. The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52:113–117, 2014.

[493] Wouter Weerkamp, Richard Berendsen, Bogomil Kovachev, Edgar Meij, Krisztian Balog, and Maarten De Rijke. People searching for people: Analysis of a people search engine log. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, pages 45–54, 2011.

[494] Rina R Wehbe, Terence Dickson, Anastasia Kuzminykh, Lennart E Nacke, and Edward Lank. Personal space in play: Physical and digital boundaries in large-display cooperative and competitive games. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2020.

[495] Jeff Weintraub. The theory and politics of the public/private distinction. *Public and private in thought and practice: Perspectives on a grand dichotomy*, 1(3), 1997.

[496] Robert S Weiss. *Learning from strangers: The art and method of qualitative interview studies*. Simon and Schuster, 1995.

[497] Laura E West. Facebook sharing: A sociolinguistic analysis of computer-mediated storytelling. *Discourse, Context & Media*, 2(1):1–13, 2013.

[498] Mark West, Rebecca Kraut, and Han Ei Chew. I'd blush if i could: closing gender divides in digital skills through education. 2019.

[499] Alan F Westin. Privacy and freedom. *Washington and Lee Law Review*, 25(1):166, 1968.

[500] Ryen W White. Skill discovery in virtual assistants. *Communications of the ACM*, 61(11):106–113, 2018.

[501] Christopher D Wickens, Justin G Hollands, Simon Banbury, and Raja Parasuraman. *Engineering psychology and human performance*. Psychology Press, 2015.

[502] Eva Wiese, Agnieszka Wykowska, Jan Zwickel, and Hermann J Müller. I see what you mean: how attentional selection is shaped by ascribing intentions to others. *PloS one*, 7(9):e45391, 2012.

[503] wikipedia. Streisand effect. Accessed: 2020-06-02.

[504] Wiktoria Wilkowska and Martina Ziefle. Privacy and data security in e-health: Requirements from the user's perspective. *Health informatics journal*, 18(3):191–201, 2012.

[505] Alex C Williams, Harmanpreet Kaur, Gloria Mark, Anne Loomis Thompson, Shamsi T Iqbal, and Jaime Teevan. Supporting workplace detachment and reattachment with conversational intelligence. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2018.

[506] Dave Wilson and Joseph S Valacich. Unpacking the privacy paradox: Irrational decision-making within the privacy calculus. 2012.

[507] Terry Winograd. Architectures for context. *Human–Computer Interaction*, 16(2-4):401–419, 2001.

[508] Diane F Witmer. Risky business: Why people feel safe in sexually expicit on-line communication. *Journal of Computer-Mediated Communication*, 2(4):JCMC245, 1997.

[509] Anne-Katrin Witte and Rüdiger Zarnekow. Transforming personal healthcare through technology-a systematic literature review of wearable sensors for medical application. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019.

[510] et al. Wong-Villacres, Marisol. Design guidelines for parent-school technologies to support the ecology of parental engagement. In *Proceedings of the 2017 Conference on Interaction Design and Children*, New York, NY, USA, 2017. ACM.

[511] Allison Woodruff. Necessary, unpleasant, and disempowering: reputation management in the internet age. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 149–158. ACM, 2014.

[512] Bin Xu, Jason Ellis, and Thomas Erickson. Attention from afar: Simulating the gazes of remote participants in hybrid meetings. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, pages 101–113. ACM, 2017.

[513] Ramin Yaghoubzadeh, Marcel Kramer, Karola Pitsch, and Stefan Kopp. Virtual agents as daily assistants for elderly or cognitively impaired people. In *International workshop on intelligent virtual agents*, pages 79–91. Springer, 2013.

[514] Lisa C Yamagata-Lynch. Understanding cultural historical activity theory. In *Activity systems analysis methods*, pages 13–26. Springer, 2010.

[515] Jie Yang, Leejay Wu, and Alex Waibel. Focus of attention in video conferencing. Technical report, CARNEGIE-MELLON UNIV PITTSBURGH PA SCHOOL OF COMPUTER SCIENCE, 1996.

[516] et al. Yarosh, Svetlana. Developing a media space for remote synchronous parent-child interaction. In *Proceedings of the 8th International Conference on Interaction Design and Children*. ACM, 2009.

[517] Svetlana Yarosh. Supporting long-distance parent-child interaction in divorced families. In *CHI'08 Extended Abstracts on Human Factors in Computing Systems*, New York, NY, USA, 2008. ACM.

[518] Svetlana Yarosh and Gregory D. Abowd. Mediated parent-child contact in work-separated families. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, New York, NY, USA, 2011. ACM.

[519] Svetlana Yarosh, Anthony Tang, Sanika Mokashi, and Gregory D. Abowd. "almost touching": Parent-child remote communication using the sharetable system. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, pages 181–192, New York, NY, USA, 2013. ACM.

[520] Weilie Yi and Dana Ballard. Recognizing behavior in hand-eye coordination patterns. *International Journal of Humanoid Robotics*, 6(03):337–359, 2009.

[521] Seounmi Youn. Teenagers' perceptions of online privacy and coping behaviors: a risk–benefit appraisal approach. *Journal of Broadcasting & Electronic Media*, 49(1):86–110, 2005.

[522] Tai-Kuei Yu, Long-Chuan Lu, and Tsai-Feng Liu. Exploring factors that influence knowledge sharing behavior via weblogs. *Computers in human behavior*, 26(1):32–41, 2010.

[523] Xuejing Yuan, Yuxuan Chen, Aohui Wang, Kai Chen, Shengzhi Zhang, Heqing Huang, and Ian M Molloy. All your alexa are belong to us: A remote voice control attack against echo. In *2018 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, Dec 2018.

[524] Beste F Yuksel, Penny Collisson, and Mary Czerwinski. Brains or beauty: How to engender trust in user-agent interactions. *ACM Transactions on Internet Technology (TOIT)*, 17(1):1–20, 2017.

[525] Aristea M Zafeiropoulou, David E Millard, Craig Webber, and Kieron O'Hara. Unpicking the privacy paradox: can structuration theory help to explain location-based privacy decisions? In *Proceedings of the 5th Annual ACM Web Science Conference*, pages 463–472. ACM, 2013.

[526] Nadia Zaheer and Peter Trkman. An information sharing theory perspective on willingness to share information in supply chains. *The International Journal of Logistics Management*, 2017.

[527] Cha Zhang, Qin Cai, Philip A Chou, Zhengyou Zhang, and Ricardo Martin-Brualla. Viewport: A distributed, immersive teleconferencing system with infrared dot pattern. *IEEE MultiMedia*, 20(1):17–27, 2013.

[528] Nan Zhang, Xianghang Mi, Xuan Feng, XiaoFeng Wang, Yuan Tian, and Feng Qian. Dangerous skills: Understanding and mitigating security risks of voice-controlled third-party functions on virtual personal assistant systems. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 1381–1396. IEEE, 2019.

[529] Xuan Zhao, Niloufar Salehi, Sasha Naranjit, Sara Alwaalan, Stephen Voida, and Dan Cosley. The many faces of facebook: Experiencing social media as performance, exhibition, and personal archive. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1–10, 2013.

[530] Martina Ziefle, Julian Halbey, and Sylvia Kowalewski. Users' willingness to share data on the internet: Perceived benefits and caveats. In *IoTBD*, pages 255–265, 2016.

[531] Martina Ziefle, Simon Himmel, and Wiktoria Wilkowska. When your living space knows what you do: Acceptance of medical home monitoring by different technologies. In *Symposium of the Austrian HCI and Usability Engineering Group*, pages 607–624. Springer, 2011.

[532] Manuela Züger, Sebastian C Müller, André N Meyer, and Thomas Fritz. Sensing interruptibility in the office: A field study on the use of biometric and computer interaction sensors. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2018.

List of author's work mentioned in this thesis but not included in Statement of Contributions

Kuzminykh, Anastasia, and Sean Rintel. Low Engagement As a Deliberate Practice of Remote Participants in Video Meetings. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1-9. 2020.
https://doi.org/10.1145/3334480.3383080

Kuzminykh, Anastasia, and Jessica R. Cauchard. Be Mine: Contextualization of Ownership Research in HCI. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1-9. 2020.
https://doi.org/10.1145/3334480.3383058

Wehbe, Rina R., Terence Dickson, Anastasia Kuzminykh, Lennart E. Nacke, and Edward Lank. Personal Space in Play: Physical and Digital Boundaries in Large-Display Cooperative and Competitive Games. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1-14. 2020.
https://doi.org/10.1145/3313831.3376319

Pietroszek, Krzysztof, Anastasia Kuzminykh, James R. Wallace, and Edward Lank. Smartcasting: a discount 3D interaction technique for public displays. In Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design (OzCHI'14), pp. 119-128. 2014.
https://doi.org/10.1145/2686612.2686629

Kuzminykh, Anastasia, and Jessica R. Cauchard. How We Own Drones: On the Sense of Ownership in the Drone Design. iHDI@ CHI. 2020.
http://ceur-ws.org/Vol-2617/paper5.pdf

Kuzminykh, Anastasia, James Wallace, Richard Zanibbi, Stacey Scott, and Edward Lank.

How Changing Communication Channels Affect Communication Patterns: Implications for the Design of Smart Objects. Technical Report, University of Waterloo, 2014.
Presented at the Workshop on Social Internet of Things (OzCHI 2014) and Workshop on Supporting "Local Remote" Collaboration (CSCW 2015).
cs.uwaterloo.ca/sites/ca.computer-science/files/uploads/files/cs-2014-26.pdf
localremote.files.wordpress.com/2015/02/presence-co-presence.pdf

Selected Mainstream Media Coverage of Thesis Contents

Parent's Information Needs in Use of Monitoring Technology (Chapter 5)

"New study from Waterloo suggests parents need to give their kids privacy."
Video news segment; CTV Kitchener. Aug, 2019
kitchener.ctvnews.ca/new-study-from-waterloo-suggests-parents-need-to-give-their-kids-privacy-1.4535721

"Kid tracking apps should avoid feeling like an 'ankle bracelet,' study says."
Paula Duhatschek, CBC News. Aug, 2019
www.cbc.ca/news/canada/kitchener-waterloo/kid-student-tracking-apps-1.5231991

"Kid tracking apps and information needs of parents."
Radio Live Interview. Kitchener today with Briane Bourke. Aug. 6, 2019
www.570news.com/audio/kitchener-today-with-brian-bourke/

"When is too much monitoring?"
Steve Kannon, The Observer. Aug. 2019
observerxtra.com/2019/08/08/when-is-too-much-monitoring/

"How much is too much? Researchers balance information needs of parents with privacy of young children."
Mirage News. July 2019
www.miragenews.com/how-much-is-too-much-researchers-balance-information-needs-of-parents-with-privacy-of-young/

Anthropomorphized Perceptions of Conversational Agents (Chapter 6)

"People Too Trusting of Virtual Assistants."
ACM Tech News. Jan.6, 2020
technews.acm.org/archives.cfm?fo=2020-01-jan/jan-06-2020.html

"The more personable a virtual assistant seems, the more data you'll share."
Nicole Bogart, CTVNews.ca. Jan.3, 2020
www.ctvnews.ca/sci-tech/the-more-personable-a-virtual-assistant-seems-the-more-data-you-ll-share-study-1.4752221

"University of Waterloo Study Found Consumers Believe Siri is Disingenuous. . ."
Bret Kinsella. Voicebot.ai. Jan.5, 2020
voicebot.ai/2020/01/05/university-of-waterloo-study-found-consumers-believe-siri-is-disingenuous-and-cunning-compared-to-alexa-as-genuine-and-caring/

"How Humanizing Virtual Assistants Earns Consumer Trust."
Laurie Sullivan, MediaPost, Jan.2, 2020
www.mediapost.com/publications/article/345201/how-humanizing-virtual-assistants-earn-consumer-tr.html

"Trusting virtual assistants could make us more revealing."
Christine Clark, KitchenerToday, Jan.3, 2020
www.kitchenertoday.com/local-news/trusting-virtual-assistants-could-make-us-more-revealing-uw-study-suggests-1995791

"People are trusting virtual assistants way too much."
Technology.org, Jan.03, 2020
www.technology.org/2020/01/03/people-are-trusting-virtual-assistants-way-too-much/

"People Too Trusting of Virtual Assistants."
Live Interviews across Canada. CBC Radio. Jan.7, 2020
(Toronto, Ottawa, Kingston-Barrie, London, Thunder Bay, Windsor, Kitchener, Sudbury)

Functional Classification of Attention in Video Meetings (Chapter 4)

"How research can enable more effective remote work"
Jaime Teevan, Microsoft Research Blog. Apr. 9, 2020
www.microsoft.com/en-us/research/blog/how-research-can-enable-more-effective-remote-work/

"Why is it so hard to pay attention in videoconferences?"
Waterloo Stories, Apr.22, 2020
uwaterloo.ca/stories/news/why-it-so-hard-pay-attention-videoconferences