Understanding Public Agencies' Expectations and Realities of AI-Driven Chatbots for Public Health Monitoring

Eunkyung Jo*
University of California
Irvine, CA, United States
eunkyuj@uci.edu

Young-Ho Kim
NAVER AI Lab
Republic of Korea
yghokim@younghokim.net

Sang-Houn Ok
NAVER Cloud
Republic of Korea
okgosu.exp@navercorp.com

Daniel A. Epstein
University of California
Irvine, CA, United States
epstein@ics.uci.edu

ABSTRACT

Advances in artificial intelligence (AI) offer the potential for chatbots to support public health monitoring by automating tasks traditionally performed by frontline workers. While introducing AI impacts public agency workers across decision-making, administration, and monitoring roles, the perceptions of workers regarding these technologies and their actual impact on labor are underexplored. We examine the case of CareCall, a large language model (LLM)-driven chatbot used to monitor socially isolated individuals, by interviewing 21 public agency workers across 13 sites involved in its adoption and rollout. We find that CareCall helped expand public reach but increased burdens on frontline workers due to insufficient resources and new labor demands, such as handling lapses in user engagement. We discuss how implementing LLM-driven chatbots in public health contexts can complicate decision-makers' articulation work and impose additional maintenance work on frontline workers. We recommend AI chatbots in this space leverage public infrastructure and incorporate fallback mechanisms.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; **Natural language interfaces**; • **Computing methodologies →** *Natural language generation.*

KEYWORDS

Chatbot, Large language models, Public sector AI, Public health monitoring, Social isolation

ACM Reference Format:
Eunkyung Jo, Young-Ho Kim, Sang-Houn Ok, and Daniel A. Epstein. 2025. Understanding Public Agencies' Expectations and Realities of AI-Driven Chatbots for Public Health Monitoring. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3706598.3713593

1 INTRODUCTION

In recent years, the significant surge in AI has permeated various public service sectors that often grapple with managing and making decisions for large populations, including child welfare [6, 9, 34–37, 65, 69], homelessness services [44], unemployment services [18, 55], and community health interventions [28, 39, 52, 81]. AI systems are often proposed to help social workers consider various factors to make fairer decisions while managing high volumes of social service requests and referrals [6, 9, 18, 36, 37, 44, 55, 65, 69] or help community health workers optimize resource allocation for large-scale health interventions [28, 39, 52, 81]. One notable area of interest is public health monitoring, which involves the continuous and systematic efforts of public agencies to collect, analyze, and interpret data to safeguard and promote the health and wellbeing of populations [74]. Public health monitoring aims to achieve various goals, including containing the spread of infectious diseases [8, 19, 49, 50, 68], promoting awareness of the public's wellbeing routines [23, 67], and improving health outcomes in underserved communities [12, 26, 53, 58, 60, 78]. Current practices of public health monitoring rely heavily on frontline workers for recurrent data collection from populations. This type of monitoring is burdensome on them and other stakeholders as public health agencies often have fewer staff or resources than needed [12, 26, 29, 59, 60, 78]. To alleviate this burden and extend reach, public agencies are increasingly considering AI technologies, such as LLM-driven chatbots, to assist or automate some of these monitoring tasks traditionally performed by frontline workers [27, 28, 30, 31].

Given the scale of public services, adopting and rolling out public sector AI systems requires collaboration among stakeholders in various roles. For example, stakeholders must decide which technology to adopt to achieve public service goals. These decisions are often driven by perceptions of the technology's capabilities and capacity to address challenges the organization faces. Day-to-day labor is also impacted as frontline work changes to accommodate the AI systems for managing the public, and administrative tasks arise to integrate the systems into complex public service infrastructures. Thus, when AI-mediated interventions are introduced, they not only change end-user interactions but also reshape the overall operation of public sectors and affect broader stakeholders. While prior work has explored stakeholder needs around AI systems in public services, it largely focused on frontline workers who directly use AI systems in their daily work [9, 27, 69, 78], limiting our understanding of the full impact of these AI technologies. Understanding how broader stakeholders in public agencies perceive the promise of these technologies and experience the real-world impact can help designers of future AI for public health monitoring to address multi-stakeholder needs and better communicate their opportunities and limitations.

To gain a more holistic understanding of how public sector AI operates within existing human infrastructures and shapes their overall institutional roles, we examine the case of CareCall, an LLM-driven chatbot developed to monitor socially isolated individuals'

*Eunkyung Jo conducted part of this work as a research intern at NAVER AI Lab.

CHI '25, April 26-May 1, 2025, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04...$15.00
https://doi.org/10.1145/3706598.3713593

health via check-up phone calls. Since its launch in November 2021, CareCall has gradually expanded to serve over 30,000 individuals across more than 140 municipalities in South Korea as of December 2024. As a rare example of an LLM-driven chatbot deployed in a real-world public health context across municipalities with varied characteristics (e.g., urban, rural) and involving workers in varying roles (e.g., frontline monitoring, decision-making, administration), CareCall provides a useful case for understanding the multi-stakeholder perspectives and experiences of the AI chatbot deployment for public health monitoring. Through the case of CareCall, we ask: **How do public agency workers across roles expect AI chatbots to assist in public health monitoring, and how do they experience the technology in reality?** Answering this question is particularly important to understand the benefits and drawbacks of AI chatbots relative to other approaches for public health monitoring, such as human-only approaches or other technologies for monitoring.

To answer this question, we interviewed 21 public agency workers involved in the adoption and rollout of CareCall across 13 sites to monitor over 4,000 individuals. Workers had varied roles in the rollout ranging from deciding whether and what technology to adopt, monitoring the data collected by the AI system to follow up with concerning cases, and taking administrative tasks to operate the system. Through the interviews, we found that public agency workers had previously struggled to regularly monitor populations needing care through phone calls or home visits (i.e., a *human approach*) due to a shortage of frontline workers. Although the agencies had tried to introduce technologies like passive sensing systems and automated voice-based systems (i.e., *hardware-dependent technologies*) to address these constraints, they created new labor demands, such as managing false alarms and maintaining devices, and their high costs still limited public reach. When CareCall was introduced, decision-makers had expectations that AI would reach more people and achieve the desirable monitoring frequency, which was largely realized. Frontline workers also valued that CareCall unexpectedly provided a window for care recipients to communicate different care needs. However, frontline and administrative workers often felt that their workload was exacerbated as the introduction of CareCall rarely involved scaling up the human resources necessary to manage the expanded care and demanded new types of labor, such as handling lapses in user engagement.

We use the framework of articulation work [71] to reflect on and discuss our findings, highlighting the changes in human labor that introducing AI chatbots requires. Decision-makers face unique challenges in conducting the required articulation work, particularly due to the open-ended nature of LLM-driven chatbots and the lack of established guidelines and best practices for these emerging technologies. We also highlight the importance of recognizing the maintenance work that AI chatbots impose on frontline workers, especially handling lapses in user engagement. Lastly, we provide implications for public agencies considering the use of AI chatbots for public health monitoring, focusing on the potential of open-ended conversations to identify unmet care needs and the need to assess the impact of AI implementation on the labor demands of their workforce. For designers and developers aiming to make AI chatbots usable for public health monitoring, we suggest opportunities to piggyback on public infrastructure to enhance scalability, incorporate fallback mechanisms to address lapses in engagement, and leverage passive sensing as a complement to AI chatbots.

Key contributions of this work include:

- An empirical understanding of the expectations and realities of public agencies in deploying AI chatbots for public health monitoring. We interviewed 21 public agency workers involved in the adoption and rollout of CareCall in varying roles, ranging from decision-making, frontline monitoring, and administration. While decision-makers' expectations that AI chatbots would expand public reach were largely met, frontline and administrative workers often experienced an increased burden due to the lack of necessary resources and new labor demands, such as handling lapses in user engagement. Despite the challenges, frontline workers saw CareCall as a tool for identifying unmet care needs.
- Insights into the articulation work required by decision-makers in implementing AI chatbots for public health monitoring, focusing on the challenges posed by the open-ended nature of LLM-driven chatbots, and the maintenance work that AI chatbots impose on frontline administrative workers to address end-user lapses in using these systems. We further offer suggestions for public agencies considering the use of AI chatbots for public health monitoring, particularly around assessing the impacts of AI implementation on labor demands, and implications for designers and developers aiming to make AI chatbots usable for public health monitoring, particularly around piggybacking on public infrastructure to enhance scalability and incorporating fallback mechanisms to address lapses in user engagement.

2 RELATED WORK

Understanding public agencies' expectations and realities of AI chatbots for public health monitoring draws on past HCI work on human infrastructures in healthcare, AI in the public sector, and technology for large-scale health monitoring.

2.1 Human Infrastructures in Healthcare

A vast and growing body of HCI and CSCW research has focused on the humans who make technology work in complex healthcare settings through the lens of human infrastructures, or organization of human labor required for collaborative work [46]. While infrastructures are traditionally understood as the physical and technological foundations of human activities (e.g., electric grids, telecommunication networks), human infrastructure is another critical aspect of understanding modern healthcare services, which necessitates collaboration among diverse professionals [70].

Introducing new technology into complex sociotechnical infrastructures inevitably requires continuous human efforts to adapt to unanticipated real-world scenarios by adjusting plans, reallocating resources, and coordinating efforts [4, 22, 66, 71]. This type of work is characterized as *articulation work*, or "a set of activities required to manage the distributed nature of cooperative work" [71]. In healthcare settings, studies have examined *maintenance work*, as a form of articulation work [66], that frontline health workers perform to anticipate, repair, and reconfigure infrastructural arrangements, thereby maintaining the community health infrastructures [64, 72, 78]. For example, Verdezoto et al. highlighted maintenance work that community health workers performed to repair the lack of social and material arrangements (e.g., limited availability of data entry personnel, a shortage of computers with internet connection) for managing survey data from populations. Extending prior work, in this study, we examine the types of articulation work and maintenance work various stakeholders perform to adopt and operate AI chatbots for public health monitoring.

2.2 AI in the Public Sector

AI technologies are increasingly being proposed as a means to overcome resource constraints in public services, particularly in health and welfare. In community health interventions, AI tools have been explored to optimize resource allocation by identifying individuals who would benefit most [28, 39, 52, 81]. In social welfare contexts, researchers have examined how AI might assist in making fairer decisions while managing high volumes of social service requests and referrals in contexts like child maltreatment screening [6, 9, 34–37, 65, 69], housing allocation for unhoused individuals [44], and job placement for unemployed individuals [18, 55]. While prior work on AI tools for public services has focused on resource allocation and decision-making support, AI can also assist in other essential public service tasks traditionally conducted by frontline workers, such as regularly collecting data from populations for public health monitoring. For example, prior work has shown how AI chatbots can be used to offload the frontline monitoring burden by automating the collection of personal health data [30, 31, 41].

While prior work has aimed to better understand the perspectives of those who directly interact with the AI systems in public services for their daily work, it has often overlooked other crucial stakeholder groups who might have substantially different needs around these tools [9, 27, 69, 78], such as decision-makers in public agencies involved in the AI adoption and other indirect users whose work has been impacted by the introduction of AI. Recently, a few studies have begun examining the perspectives of workers involved in AI-mediated decisions in varying roles (e.g., supervisors, agency leaders) [34–37]. For instance, Kawakami et al. found that frontline workers had different target outcomes for child maltreatment screening than the AI tools but faced organizational pressures to disagree with the algorithmic decisions [36], suggesting the need to understand the viewpoints of those who hold higher power and responsibility to shape the adoption and rollout of AI tools in the public sector. In response to this call, a recent study engaged with decision-makers in public agencies to understand their views on adopting new AI tools, highlighting the differing perspectives of decision-makers and frontline workers on the validity and value of these tools [34]. We extend prior work by examining perspectives from public agency workers involved in the adoption and rollout of AI chatbots in varying roles and how their needs align and differ in the context of public health monitoring.

2.3 Technology for Large-Scale Health Monitoring

Monitoring the health and wellbeing of large populations demands significant time and effort from public health agencies to conduct recurrent data collection [12, 26, 29, 59, 60, 78]. Prior work in the HCI and CSCW communities has often studied or proposed technologies, such as chatbots and mobile apps, to support large-scale health monitoring in contexts like contact tracing [8, 41, 49, 50], maternal and child health education [12, 26, 29, 33, 59, 60, 63, 78], and social isolation intervention [30, 31], mainly focusing on supporting frontline workers who manually collect health-related data from populations. Meanwhile, studies in the Ubiquitous Computing community have proposed more technical approaches to large-scale health monitoring through sensor technologies. Studies have frequently proposed sensor-based in-home monitoring as a mechanism to monitor various health indicators such as air quality [17, 25, 56, 82], water usage [20, 21, 75], and electricity consumption [16, 42, 61], particularly highlighting the benefits of piggybacking on public infrastructures in improving the scalability of such technologies [21, 42, 75].

Although technologies are often developed with the intention of offloading the monitoring burden, they oftentimes bring about unintended consequences. Studies have shown that technology designed to support community health work often increases the strain on frontline workers by introducing additional responsibilities [29, 38, 59, 72, 78]. Operating these systems frequently requires substantial efforts from frontline workers to go beyond the job descriptions, but such efforts tend to be unacknowledged by other stakeholders [53, 54, 62, 72, 78]. Research has characterized such overlooked contributions of frontline workers as *"invisible work"* [10]—labor that is essential to their job but unnoticed, unacknowledged, or undervalued by other stakeholders [38, 53, 54, 59, 72, 78]. For example, in their study of data-driven technologies in a long-term care facility, Sun et al. revealed that frontline workers performed substantial data work to address the breakdowns in the data infrastructure—such as repairing incorrect or incomplete data collected through sensor technologies—but such work was largely neglected by other stakeholders under the guise of innovation [72]. Findings from previous studies suggest the need for an in-depth understanding of the full extent of the labor that stakeholders perform when introducing technology to the public sector. Our work adds to this literature by specifically unpacking the expectations public agency workers have toward emerging AI technology like CareCall and how such technology actually impacts their labor in real-world settings.

3 STUDY CONTEXT: CLOVA CARECALL

We present CareCall as an example of an AI chatbot deployed to support public health monitoring. This section draws from AI and NLP literature on the underlying technology ([2, 40]), design documents written by the designers and developers of CareCall, and interviews with public agency workers.

3.1 Motivation and Design of CareCall

CareCall is an LLM-driven voice chatbot designed to support socially isolated individuals via scheduled phone calls [7]. The system's motivation stems from the recent act on the Prevention and Management of Lonely Death in South Korea [43]. This act highlights the importance of identifying and supporting individuals at risk of lonely deaths—where someone lives alone, is cut off from family or relatives, dies alone at home due to suicide or illness, and their body is found only after a delay. As South Korea has seen a sharp increase in lonely deaths over the past few years, from 2,412 in 2017 to 3,378 in 2022, societal consensus has emerged on the need to prevent these deaths [80].

Designed as an open-ended dialogue system powered by an LLM called HyperCLOVA [40], CareCall mimics the conversational style of social workers casually checking in on their care recipients. Equipped with long-term memory [1], CareCall can remember information mentioned by users and incorporate that knowledge when generating responses in future sessions. For example, if a user mentions regularly seeing a doctor for leg pain, CareCall generates a summary of this information and includes it in the model input for subsequent sessions. This allows the system to ask follow-up questions, such as, *"How is your leg feeling?"* Prior work has examined the perceptions of the design of CareCall, showing that its open-ended conversations helped mitigate the loneliness of end-users and provided public health workers with a holistic understanding of individuals [30]. Further, the long-term memory feature particularly enhanced health disclosure and fostered positive perceptions of the chatbot by offering familiarity [31]. More technical details of the system architecture can be found in prior work [1, 30, 31, 40].

3.2 Deployment of CareCall

Initially launched in Busan, Korea, in November 2021, CareCall had gradually expanded to monitor over 30,000 individuals as of December 2024 through over 140 public agencies in South Korea. CareCall was adopted and deployed by various types of public agencies that were taking care of socially isolated individuals, including local governments, community health centers, and Veterans Affairs offices. The scale of CareCall deployments varied significantly, ranging from fewer than ten to several thousand individuals, depending on the reach of the public agencies involved.

Each public agency had slightly different criteria for the target users of CareCall in terms of the age group or specific health conditions, though they all shared the overarching characteristic of monitoring socially isolated people. Most public agencies deployed CareCall to middle-aged (40s to 60s) and older adults (60s or older) living alone who were at a low socioeconomic status (e.g., below 50% of median household income). However, some agencies in public healthcare contexts deployed CareCall specifically to older adults with mild cognitive impairment or depression. In most cases, public officers who provide social services in neighborhoods recommended these individuals to use CareCall.

The adoption and rollout of CareCall required public agencies to undertake different tasks. Workers can be categorized as having one or more of three roles: *decision-making*, *administration*, and *frontline monitoring*. Similar roles have been described in other studies of public service technologies (e.g., decision-making [34, 36, 65], administration [36, 65], and frontline work [9, 27, 30, 32, 37, 78]). Before adopting and rolling out the system, public agencies had to make various decisions, including assessing whether this technology is suitable for achieving their public service goals, planning budgets, assigning tasks to subordinate agencies to manage the operation of the system, and developing monitoring protocols (**decision-making**). After CareCall started rolling out, frontline workers regularly monitored the call logs to see if any negative health signals were detected (*e.g.*, skipping meals, poor sleep, health issues) or if the person did not answer several calls in a row (**frontline monitoring**). When any health concerns or consecutive missed calls were detected from the call logs, the frontline workers were notified to check with the person to see if everything was okay. If there was no response to these manual calls, they either visited the individual's home or escalated the matter to local public officers. Similarly, if social service or healthcare needs were identified during manual calls, they either directly connected the individuals to those services or wrote a report to escalate the issue to local public officers. Public agencies also took on various administrative tasks to coordinate among agencies at different levels, such as compiling the list of care recipients who needed manual check-ins from frontline workers and relaying the list to local public officers (**administration**).

Given the varied scales of CareCall deployments, each agency distributed roles in different ways. For example, some upper-level local governments (e.g., provincial or city governments) deployed CareCall on a relatively larger scale and assigned administrative and monitoring roles to their subordinate institutions (e.g., neighborhood community centers or local social service agencies), involving a few to around twenty workers in deploying the system in each site. In contrast, at agencies that deployed CareCall on a smaller scale or were low on resources, one or two personnel undertook multiple roles needed for the entire process of adoption and rollout.

4 METHODS

To understand public agencies' expectations and realities of deploying AI chatbots for public health monitoring, we interviewed 21 public agency workers involved in the adoption and rollout of CareCall in varying roles, such as decision-making, administration, and monitoring. Our interview study was classified as exempt by our University's Institutional Review Board as the methodology did not involve more than minimal risk to participants.

4.1 Interview Process

In Fall 2023, we conducted individual semi-structured interviews with 17 of our 21 participants, while the remaining four participated in pairs with a colleague from the same agency (19 sessions in total). Interview sessions were conducted via conference calls (eight), phone calls (four), or in-person (seven) based on participant preference, each lasting 40-60 minutes. During the interviews, we asked about (1) their prior experiences monitoring health and wellbeing of populations using human and technological approaches, (2) the motivation for CareCall adoption and influencing factors, (3) the impact of integrating CareCall into their workflows and public health infrastructure. We offered all participants 60,000 KRW (roughly 45 USD at the time of the interviews) for their time. However, eight participants opted out of receiving study compensation to avoid the complex process of reporting external income as government officials, while the remaining thirteen were compensated.

4.2 Participants

We recruited participants by distributing flyers to public agencies deploying CareCall (14) and via snowball sampling (7), consisting of 12 females and 9 males aged from 25 to 45 (Table 2). Eligibility criteria included individuals involved in the adoption and rollout of CareCall as part of their work for at least three months.

Table 1: Information on the sites where participants were involved in the adoption and deployment of CareCall, including the scale of deployment (number of CareCall users), geographical characteristics, public service context, and target users

Site	Deployment Scale	Geographical Characteristics	Public Service Context	Target Users
SiteA	500	Mostly rural	Social welfare	Low-SES middle-aged and older adults living alone
SiteB	1,000	Mostly urban	Social welfare	Low-SES middle-aged and older adults living alone
SiteC	100	Urban, suburban	Public healthcare	Older adults living in isolated islands
SiteD	5	Rural, suburban	Public healthcare	Older adults with depression
SiteE	180	Urban	Public healthcare	Older adults with mild cognitive impairment
SiteF	40	Rural, suburban	Social welfare	Low-SES older adults living alone
SiteG	50	Rural, suburban	Social welfare	Low-SES middle-aged adults with chronic conditions
SiteH	10	Suburban	Social welfare	Low-SES middle-aged adults living alone
SiteI	30	Rural	Social welfare	Low-SES older adults living alone
SiteJ	300	Urban	Social welfare	Low-SES older adults living alone
SiteK	1,500	Urban, suburban, rural	Social welfare	Low-SES middle-aged and older adults living alone
SiteL	50	Urban, suburban	Tech incubation	Low-SES older adults living alone
SiteM	270	Urban	Tech incubation	Low-SES older adults living alone

The workers we interviewed were from 13 different sites (SiteA-M) that deployed CareCall to achieve different goals depending on their public service contexts, such as social welfare, public healthcare, and technology incubation for public services (Table 1). For example, social service agencies deployed CareCall to monitor general health and well-being of low-SES middle-aged and/or older adults living alone, aiming to prevent lonely deaths. Conversely, community health centers adopted CareCall to monitor more specific health concerns, such as depression, mild cognitive impairment, or limited healthcare access (e.g., living in an isolated island). The scale of the deployment in these sites greatly varied, ranging from five to 1,500 individuals (Table 1). The geographical characteristics of these sites varied as well, including predominantly rural, suburban, and urban areas, as well as some with a mix of these elements (Table 1). Such a classification is based on factors such as population density, urban development, and regional characteristics [48].

Depending on their resources and the scale of deployments, most of these workers were taking multiple roles in the adoption and rollout of CareCall. Ten workers participated in decision-making around the adoption and rollout of CareCall, 20 performed administrative roles, and 14 conducted frontline monitoring work (Table 2). Some of these workers were working with each other across institutions to manage CareCall deployments collaboratively. For example, P-SiteA-1 was in a provincial government and distributed some of the administration and monitoring tasks to their subordinate agencies, such as city governments (where P-SiteA-2 and P-SiteA-3 were based) and a social service agency (where P-SiteA-4 was based). Similarly, P-SiteB-1 made the decisions about the CareCall adoption in a provincial government and assigned some of the administration and monitoring tasks to their subordinate agencies, including a community care center (where P-SiteB-3 was based) and a social service agency (where P-SiteB-2 was based). We also interviewed individuals who were working as teams managing CareCall deployments within the same institutions, such as P-SiteC-1 and P-SiteC-2, P-SiteD-1 and P-SiteD-2, and P-SiteE-1 and P-SiteE-2.

Workers had varying length of involvement in the adoption and deployments of CareCall, ranging from 3-6 months (six), 6-12 months (eight), and over 12 months (seven). Some participants had prior experiences with adopting technologies other than CareCall for public health monitoring. Six of them had experience in leveraging passive monitoring systems in the context of social isolation intervention, such as movement monitors (P-SiteB-1, P-SiteA-2, P-SiteA-4), call history monitoring systems (P-SiteI), and smart plugs tracking power usage (P-SiteH, P-SiteG, P-SiteI). Four workers had experience in deploying automated voice-based systems, such as answering machines (P-SiteA-2), rule-based chatbots (P-SiteF), and social robots (P-SiteD-1, P-SiteD-2) in similar contexts.

Table 2: Participant demographics, including age, gender, and role in CareCall deployment. ID denotes their affiliated site.

ID (Gender, Age)	Role in CareCall Deployment		
	Decision-making	Administration	Frontline Monitoring
P-SiteA-1 (M, 42)		✓	
P-SiteA-2 (M, 35)		✓	
P-SiteA-3 (F, 43)		✓	
P-SiteA-4 (F, 41)		✓	✓
P-SiteB-1 (M, 35)	✓	✓	
P-SiteB-2 (F, 25)		✓	
P-SiteB-3 (F, 40)			✓
P-SiteC-1 (M, 42)		✓	✓
P-SiteC-2 (M, 41)	✓	✓	✓
P-SiteD-1 (M, 31)	✓	✓	✓
P-SiteD-2 (F, 35)		✓	✓
P-SiteE-1 (F, 28)	✓	✓	✓
P-SiteE-2 (F, 35)		✓	✓
P-SiteF (F, 43)	✓	✓	✓
P-SiteG (F, 34)	✓	✓	
P-SiteH (F, 30)	✓	✓	✓
P-SiteI (M, 38)	✓	✓	✓
P-SiteJ (F, 33)		✓	✓
P-SiteK (F, 35)		✓	✓
P-SiteL (M, 45)	✓	✓	✓
P-SiteM (M, 35)	✓	✓	✓

4.3 Data Analysis

We audio-recorded and auto-transcribed all interview sessions, manually correcting the automatic speech recognition errors in the transcripts later. All data were originally captured in Korean and were translated into English during the analysis process by the first author, who is a native Korean and fluent in English. We paraphrased some idioms and phrasings to sound more natural in English and cross-checked the validity of the translation with the second author, who is also a native Korean and fluent in English.

We used inductive thematic analysis, characterized by the generation and constant comparison of open codes to reveal underlying themes [5], to qualitatively analyze the interview transcripts. The first author open-coded the interview transcripts and revised overarching patterns and themes through several rounds of peer debriefing meetings. From this coding, we surfaced the main theme around the expectations and realities of deploying AI-driven chatbots for public health monitoring, using this to organize our results. The final codebook contained nine parent codes, such as expecting AI to help expand care and AI increasing frontline workload in practice, and 21 child codes. A central theme in our interviews was the varied forms of human labor needed by each stakeholder to manage the deployment of CareCall. This led us to use the framework of articulation work [71] to guide the discussion of our findings. Our use of articulation work was bottom-up, informed by our findings rather than presupposed ahead of analysis.

5 FINDINGS

In this section, we organize our findings by first introducing the prior experiences of agency workers in taking a human approach and using hardware-dependent technologies for public health monitoring. We then compare these experiences against the expectations and realities of using AI-driven chatbots (e.g., CareCall) for this task. Municipalities tended to introduce new technology sequentially given resource constraints and evolving policy priorities.

5.1 Prior Experiences with Human Approaches

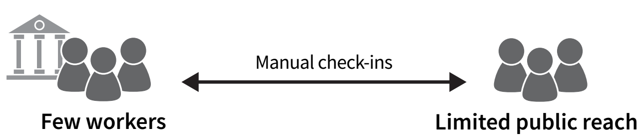

Figure 1: Human approaches—Public agencies traditionally monitored socially isolated individuals through manual check-ins, such as phone calls or home visits. With few frontline workers assigned to these tasks, only a limited number of people received regular monitoring.

Public agencies have predominantly relied on a human approach for social isolation interventions, such as phone calls or home visits. However, with a shortage of frontline workers relative to the population needing care, this approach was often perceived as overburdening. P-SiteA-2, as an administrator, acknowledged the lack of resources given to frontline workers: *"In our city, a single social worker handles over 100 socially isolated individuals, which isn't realistically manageable."* Frontline workers described similar feelings toward a human approach. P-SiteF said: *"Checking in with elderly individuals is quite a burden. I'm responsible for about 1,000, making it unrealistic to check in with everyone."* P-SiteJ echoed this, stating: *"I'm in charge of hundreds of individuals, which means that I need to call dozens of people per day. It's too time-consuming and exhausting."* Both decision-makers and frontline workers felt the resource-intensive nature of the human approach often resulted in inadequate monitoring frequency and coverage of populations in need of regular monitoring (Figure 1).

5.1.1 Insufficient Monitoring Frequency. Due to resource constraints, all workers were concerned that a human approach did not enable them to engage in the monitoring tasks as frequently as they desired. The frequency that the frontline workers manually checked in with socially isolated individuals varied greatly, ranging from yearly to weekly check-ins, but they often felt more frequent monitoring was necessary to ensure the safety of the populations. P-SiteH, a frontline social worker in charge of monitoring bed-bound elderly people, said: *"I haven't been able to check in with them as frequently as I wanted due to resource constraints."* Administrators and frontline workers often attributed the gap between the desired and actual frequency of check-ins to the resources required for home visits. P-SiteD-2 described the difficulties of relying on home visits for monitoring in a rural region: *"In big cities, people typically visit a mental health center themselves, so it's much easier for social workers to check in regularly. But in rural towns like ours, elderly people can't visit us because our center is far from their homes and public transport is poor, so we need to visit their homes. But with limited resources, it is only one or two times a month, which is obviously not enough."*

5.1.2 Limited Public Reach. All workers perceived that the resource-intensive nature of human approaches limited their ability to adequately cover populations in need of regular health monitoring. They frequently indicated that there were far more individuals who would benefit from monitoring, but only a small percentage of them were regularly monitored as they had to selectively prioritize those who were at greater health risks. P-SiteE-1, a social worker in charge of individuals who are at risk of dementia at a community mental health center, felt that they could not adequately cover the population through a human approach: *"Thousands of individuals with mild cognitive impairment are registered in our pool. We can't really check in with every one of them, so a lot of them have been unattended."* P-SiteL described a similar situation for monitoring socially isolated elderly people: *"In our city, over 25,000 older adults are living alone. Among them, those who are low-SES or have severe disabilities have social workers visiting them regularly for check-ins. However, many others are not receiving any regular check-ins."*

5.2 Prior Experiences with Dedicated Hardware

Recognizing that a human approach did not enable them to monitor as frequently or as much of the public as they would have liked, decision-makers in government agencies often adopted some hardware to monitor socially isolated populations. As mentioned in Section 4.2, several workers had experiences in leveraging passive monitoring systems in social isolation interventions—such as motion sensors, smart plugs that track power usage and ambient light levels, or a call history monitoring system—which were designed to notify frontline workers when detecting unusual patterns that indicate potential medical emergencies. In addition, a few workers had experience in deploying automated voice-based systems—such as an answering machine, a rule-based chatbot, and a social robot for social isolation intervention. Frontline workers mentioned some instances where they found such technologies beneficial. For example, P-SiteH described how the smart plug helped discover a case of a lonely death: *"There was a person who was bedridden. He would leave the TV on the whole day and turn it off when going to bed. But one day, his power usage suddenly showed up as zero, so we tried reaching out and found he had passed away."*

Decision-makers generally found that dedicated hardware helped them slightly scale up monitoring beyond human approaches. However, frontline workers felt that such hardware-dependent technologies minimally alleviated the monitoring burden as they introduced new labor demands. Further, the high cost of hardware devices limited the reach of these approaches (Figure 2).

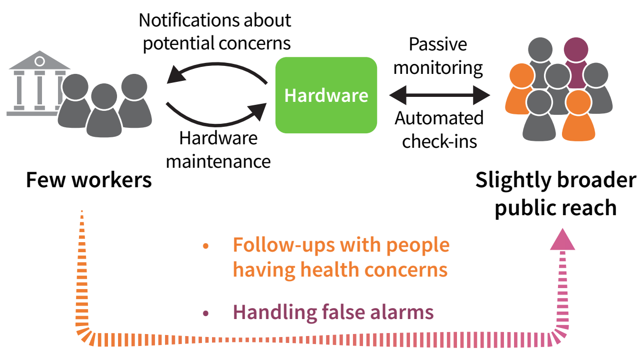

Figure 2: Hardware-dependent technologies—Decision-makers in public agencies adopted hardware for passive monitoring or automated check-ins to notify frontline workers to follow up with people having health concerns (highlighted in orange). However, frontline and administrative workers found these technologies minimally alleviated the monitoring burden, as they introduced new tasks—such as handling false alarms (highlighted in purple) and maintaining hardware, and the high costs only marginally expanded public reach compared to human approaches.

5.2.1 Introduction of New Types of Labor for Frontline Workers.

Although sensor-based systems were introduced to assist in public health monitoring, administrators and frontline workers often found these systems added new labor demands instead of reducing their burden by frequently triggering false alarms and requiring constant follow-ups. Reflecting on her experience using smart plugs to monitor socially isolated people, P-SiteG noted that the sensors were error-prone, creating a significant burden: *"Since smart plugs measure changes in the environment, like ambient light levels, we got too many false alarms when a person forgot to turn off the light before going out or something. When we got those notifications, we were supposed to check if the person was okay immediately, even if it was evening or weekend, which was quite a burden."*

In addition, the hardware dependency of these systems led frontline workers to take on additional tasks for managing the devices. P-SiteG described the hardware maintenance tasks that the smart plugs required: *"We had a lot of difficulties managing the smart plugs, as the devices would sometimes break down or participants might lose them. We also had to get the devices back when participants dropped out to install them in new participants' homes, so it was a lot of work."* Similarly, while managing a social isolation intervention for older adults with depression using social robots, P-SiteD-1 faced administrative challenges in troubleshooting the devices for the participants: *"We have to do pretty much everything for the participants in terms of setting it up and troubleshooting. They often call us and say, 'It won't turn on,' so I would drive 40 minutes to their homes to check it out. Usually, it's just a temporary issue with their WiFi or something, so I would just restart the router and come back. It's definitely taking a lot of resources to maintain the robots."*

5.2.2 Limits to Scaling Up Monitoring.

Decision-makers further perceived that the high cost of the devices constrained their ability to extend their monitoring reach to desired levels. P-SiteB-1 highlighted the financial barriers in providing care through motion sensors: *"We had to be quite selective when deploying the movement monitors because of the budget limit. There's still a long waitlist of people who want this service but haven't received it."* These costs motivated P-SiteB-1 to *"additionally adopt CareCall because it allowed us to reach a lot more individuals within the budget limit."* Other decision-makers who had experience or considered deploying smart speakers for older adults living alone also found that the high cost of the devices significantly limited their ability to scale up the intervention. P-SiteH explained that the main reason that their locale decided not to adopt smart speakers for social isolation intervention was the cost of the devices: *"There are plenty of smart speakers out there that can help monitor populations, but those are quite pricey. This limits our ability to expand care. I think budget is the primary factor when deciding which technology to adopt."*

5.3 Expectations for AI-Driven Chatbots

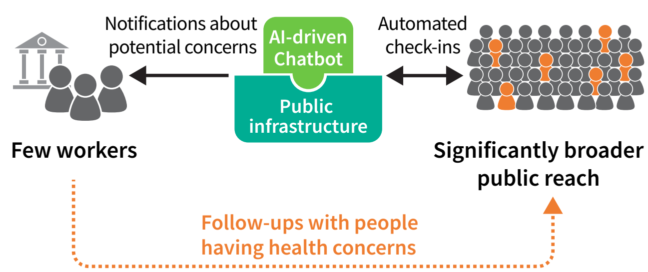

Figure 3: Expectations for AI-driven chatbots—When adopting CareCall, decision-makers in public agencies expected this AI chatbot to expand care to a much larger population through automated check-ins. They anticipated that the existing workforce could manage this expansion of care, as it would require follow-ups with only a small number of people who indicate health concerns (highlighted in orange).

When adopting CareCall, decision-makers had expectations for the AI chatbot to overcome the aforementioned limitations of a human approach and hardware-dependent technologies, enabling more frequent monitoring for a larger number of people while alleviating the burden on workers (Figure 3). They decided to adopt the technology because they thought it would help scale up monitoring reach and frequency without needing to scale up workers. Decision-makers expected that they would be able to monitor significantly more of the public, effectively simulating human calls, at a lower cost than human approaches or passive monitoring.

5.3.1 Expansion of Public Reach with Increased Monitoring Frequency. We observed that decision-makers saw CareCall as a means to expand public reach through automated monitoring. They believed technology, particularly AI systems like CareCall, would help overcome resource constraints in public health monitoring. P-SiteE-2 sought technological solutions to expand care amidst resource constraints in community health: *"The number of people at risk of dementia has skyrocketed in our municipality, but our resources haven't really increased much. We figured we would need technology to more efficiently provide care to those people."* P-SiteA-1, who administered the rollout of CareCall, similarly believed AI would help overcome resource constraints and scale up care for socially isolated individuals: *"There are a lot of people who would benefit from care but haven't gotten any because of our limited resources. With AI, we hope to provide care to more people."*

As mentioned in Section 5.1, many decision-makers believed more frequent monitoring was necessary to ensure the safety of these populations than what a human approach could achieve. Decision-makers and administrators envisioned that CareCall, as an AI chatbot performing automated check-up calls, would significantly increase monitoring frequency. P-SiteC-1 believed that CareCall's automated monitoring would achieve the desired frequency of check-ins for isolated island residents: *"Our team has been visiting small, isolated islands to offer free medical check-ups for the past few years. But we only get to visit each island once or twice a year, so it wasn't really proper monitoring. With CareCall, we can now check in with them weekly."* Managing socially isolated older adults in a rural town, P-SiteA-2 also perceived CareCall would allow them to monitor more frequently: *"We used to visit each individual in our pool monthly at most because of resource constraints. Now we're able to check in with them more often with CareCall."*

5.3.2 Requiring Minimal Human Labor. When introducing CareCall, decision-makers further viewed it as an efficient AI tool that could offload the burden on frontline workers. As human approaches to monitoring tasks were often perceived as overburdening for frontline workers, decision-makers explicitly aimed to offload their burden through AI adoption. P-SiteG described: *"Our aim of adopting CareCall was to offload the frontline workers' burden. They've been asked to check in with people who are at risk of lonely deaths at least once or twice a week, but we hoped CareCall could check in on their behalf and ease their workload."* Frontline workers had similar hopes, such as P-SiteJ: *"Our city viewed that AI would offload our burden by automatically checking in with individuals on my behalf, allowing me to focus on those are in greater need of care."*

Frontline workers believed that the AI system would require minimal human labor in frontline monitoring as it automates repetitive inquiries with the public and generates logs for review. P-SiteC-1 stated: *"I think the biggest strength of CareCall is that it saves time. AI automatically calls people regularly and asks different questions so that we can see whether they have issues in different aspects of health through the call logs."* P-SiteD-1 similarly mentioned: *"Before using CareCall, the only way that we could monitor how they were doing was by giving them a call or visiting their home. With CareCall, I would no longer need to do it manually; instead, I can see whether and what health issues they have through the call logs, so I can only follow up with people who need my attention."*

5.4 Realities of AI-Driven Chatbots

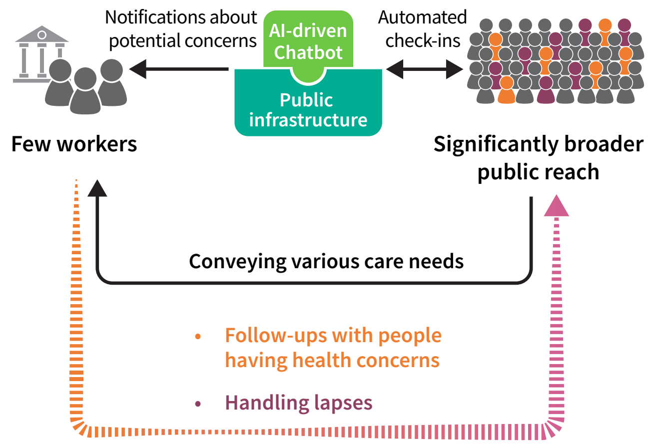

Figure 4: Realities of AI-driven chatbots—The introduction of CareCall fulfilled decision-makers' expectation to expand public reach, particularly due to piggybacking on public infrastructure, while also serving as a channel for individuals to communicate care needs. However, frontline and administrative workers felt that their workload increased as the expansion did not involve scaling up staff. Using AI chatbots also introduced new labor demands, requiring frontline and administrative workers to follow up not only with people with health concerns (highlighted in orange) but also with those who lapsed in engaging with this chatbot intervention (highlighted in purple).

When CareCall was introduced, decision-makers' expectations of expanding public reach and increasing monitoring frequency were largely fulfilled, as CareCall was indeed able to reach a broader audience than human approaches or dedicated hardware. However, all workers' expectations of requiring minimal additional human labor were unmet. Frontline workers often felt the introduction of CareCall exacerbated their workload by expanding care without necessary resources and demanding new types of labor. Meanwhile, frontline workers noted an unexpected benefit, as CareCall provided a window to communicate different care needs (Figure 4).

5.4.1 Expansion of Public Reach with Increased Monitoring Frequency. In many municipalities, decision-makers felt that CareCall successfully expanded public reach and increased monitoring frequency through automated check-up calls. Decision-makers in public agencies introduced CareCall primarily to scale up care to underserved populations. P-SiteB-1 described: *"In our province, CareCall users are mostly individuals whom the frontline workers* *wanted to check in on but couldn't because they are at relatively lower risk."* P-SiteI similarly illustrated how CareCall helped scale up care: *"We've been visiting elderly veterans for check-ins, but we could only visit up to ten homes per day. We figured it would be nice to reach out to more people with the help of technology. With CareCall, we're able to monitor around 100 individuals per day."*

All workers noted that CareCall enabled more frequent monitoring of care recipients. P-SiteE-1 valued that CareCall allowed weekly check-ins with individuals with mild cognitive impairment: *"We have thousands of people in our pool, so regular check-ins had been difficult. With CareCall, we can now check in more frequently."* Similarly, P-SiteC-1 valued that CareCall's automated monitoring allowed weekly check-ins with isolated island residents: *"CareCall lets us check in weekly. I love that the system automatically checks in with people and flags high-risk cases that need our attention."*

Decision-makers highlighted that one factor that enabled the expansion of public reach was CareCall's reliance on public infrastructure, specifically telephone networks. Unlike hardware-dependent technologies, CareCall offers chatbot interactions through scheduled phone calls, eliminating the need for dedicated devices and reducing adoption and maintenance costs. Decision-makers and administrators perceived that this reliance on public infrastructure as key to scaling care within budget constraints, often comparing it to passive monitoring systems needing dedicated hardware. P-SiteA-2 emphasized how CareCall's low cost helped expand its reach: *"Sensor-based systems require substantial initial costs, so it's difficult to scale up. With $10,000 budget, we can reach around 1,000 people through CareCall, but only dozens with those sensor-based systems."* Decision-makers also compared the budget required for deploying CareCall to the one required for deploying smart speakers, another popular type of technology used in elderly care. P-SiteE-1, a social worker responsible for people with mild cognitive impairment, highlighted that CareCall significantly increased the number of people that they could reach within the same budget compared to deploying smart speakers: *"Our team is also deploying smart speakers for older adults so that they can play music, etc. With the limited budget, we can only reach ten people, whereas CareCall allows us to reach around 180 individuals with the same budget."* P-SiteH similarly valued that CareCall's independence of hardware reduced costs compared to smart speakers: *"Smart speakers are potentially helpful for monitoring populations, but they are expensive. We wouldn't be able to provide care to as many people as we want to. CareCall didn't require purchasing devices, which allowed us to provide care to more individuals."*

Although CareCall overall reduced costs for monitoring individuals, some decision-makers deployed CareCall in conjunction with hardware-dependent approaches to better align with how the public was interested in engaging with technology. P-SiteK monitored some individuals in their municipality with CareCall, and others with a smart plug, primarily for those resistant to using a chatbot but still needed monitoring: *"We realized that some people have reservations about talking to AI. We mostly deployed CareCall for those comfortable with AI, while using smart plugs for individuals who didn't like interacting with AI but still needed continuous monitoring due to their poor health."* P-SiteI similarly deployed a smart plug alongside CareCall, noting that passive monitoring was better for some of their municipality because of other health conditions: *"Many older adults experience hearing loss, so it's difficult for them to engage in proper phone calls. For those people, we installed the smart plugs to monitor power usage, while deploying CareCall for people with good hearing who needed some emotional support."*

5.4.2 Expected Expansion of Care without Necessary Resources. Although CareCall was introduced to expand care, in reality, frontline workers felt that it was rarely accompanied by the necessary resources to support its implementation. Decision-makers often viewed AI systems like CareCall as highly efficient and requiring minimal human labor. As a result, even when the AI tool was introduced to expand public reach and increase monitoring frequency, they did not hire more staff members to manage the expanded care in most cases, leaving existing frontline workers and administrators to take on extra tasks. For example, with the introduction of CareCall, P-SiteK was assigned to monitor logs from around 1,500 people on her own, which she felt was unmanageable: *"Our province set the goal to provide the service to 2,000 people from the beginning. Currently, I'm the only person monitoring CareCall logs in our province, and I also have other responsibilities. I can't really keep up with monitoring that many people."*

Frontline workers and administrators described how decision-makers' perceptions of AI might have impacted resource allocation for AI interventions. P-SiteC-1, a frontline worker at a medical center, shared his perspective on the possible rationale behind the insufficient resource allocation: *"Decision-makers probably want AI to take care of 1,000 people and have frontline workers focus on 10% at higher risk. They wouldn't want to hire more people when introducing AI because it's supposed to be efficient. Introducing AI is not to allocate more resources."* P-SiteB-1, who made decisions about the deployment of CareCall in their province, acknowledged they underestimated the resources required to manage CareCall: *"We originally planned on deploying the system with 2,000 individuals but had to cut down to 1,000 when we saw how much it disrupted their other responsibilities. Like, they could barely keep up with their primary tasks because they had to spend almost two full days every week just for CareCall monitoring."*

Frontline workers often felt that the monitoring tasks introduced by CareCall substantially distracted them from other responsibilities. P-SiteB-3's primary role as a frontline social worker was to assign home care aides to bedridden elderly individuals. However, with CareCall's introduction, the provincial government tasked her with monitoring logs from 600 people—an entirely new and additional responsibility beyond her existing duties. P-SiteB-3 described the impact of CareCall on her frontline workload: *"I recognize how CareCall can be helpful in some cases. I would have loved it if checking in with older adults was my day job because it can check in with hundreds of people in just a few hours, a task that would have taken me a whole week. But the thing is, my day job is sending home care aides. For me, CareCall added a whole new set of tasks on top of it."*

Frontline workers thus highlighted the need to allocate additional resources to handle the expanded care with CareCall. For example, P-SiteA-3 stated: *"When public agencies start new projects, existing frontline workers always end up with additional tasks, even though we already deal with heavy workloads. There should be dedicated staff that can focus on CareCall monitoring."* All workers argued that introducing AI would only increase their burden without extra resources, as it inevitably requires human labor. P-SiteG stated: *"The aim was to offload public officials' burden, but if the project is managed by existing workers, it only adds to our burden. AI does the monitoring, but follow-ups need human labor."* In this climate, a couple of government agencies began subcontracting CareCall monitoring tasks. P-SiteB-1, a decision-maker at a provincial government, described: *"We're planning to make budgets to hire people solely for CareCall monitoring next year. Without this, our social workers can't keep up with their other responsibilities."*

5.4.3 Introduction of New Types of Labor for Frontline and Administrative Workers. As mentioned in Section 5.3.2, many decision-makers hoped CareCall would reduce the burden on frontline workers by automating repetitive check-up calls. When introduced to real-world settings, some administrators found it helpful for reducing the need for hardware maintenance, as it piggybacked on public infrastructure, unlike prior hardware-dependent technologies. P-SiteG valued that CareCall did not require any device maintenance—including handling breakdowns and device losses, installing devices for new participants, and retrieving devices from dropouts—unlike smart plugs that they had previously handled the administration and rollout of: *"Managing smart plugs was a lot of work. CareCall doesn't really require any of that."* Similarly, P-SiteI highlighted the convenience of deploying CareCall in comparison to administering social robots: *"Those robots need to be charged regularly and would occasionally have technical issues. I like that CareCall requires minimal effort in managing the system."*

However, frontline workers felt that CareCall's introduction overall exacerbated their burden by demanding new types of labor. As mentioned in Section 3, CareCall was designed to notify officials in charge of individuals who eventually did not answer after three call attempts, with each government having its own protocols for handling such lapses. For example, SiteB had 'the same-day protocol,' requiring frontline workers to check in on every individual who missed CareCall calls the same day. Although the goal was to ensure the safety of care recipients, frontline workers often found such protocols overburdening. P-SiteB-3, a frontline worker assigned to monitor 1,000 people after CareCall's introduction, said she had to spend a full day or two every week just for follow-ups on missed calls: *"Every Wednesday, I have to mentally prepare myself before checking the call status. We never know the turnout, so I get a bit anxious every week. Even though CareCall tries three times, about 10% of the users don't answer, so I have to call around 100 people on average every Wednesday."* P-SiteF also perceived adhering to the protocols for handling lapses was challenging: *"We get anxious when the participants don't answer calls. If I can't confirm the person is doing okay through CareCall, human calls, or visits, we have to call 911 to forcibly open their front door. It would be a major inconvenience for them, so I'd rather not have to do it."* Administrative workers also reported that handling lapses required them to undertake additional tasks to coordinate between multiple institutions. P-SiteB-2 described the complex coordination process of addressing lapses: *"The community care center lists individuals needing public officers' attention after monitoring the call logs and relays the list to us. We then relay it to the Provincial government so they can reach out to officers in each neighborhood. The whole process is very complex."* These administrative tasks for handling lapses were perceived as time-consuming and burdensome by other workers as well.

Frontline workers viewed handling lapses in user engagement with CareCall as redundant, as these were not typically due to health concerns. P-SiteC-2, responsible for isolated island residents, thought lapses often occurred because participants disliked AI: *"CareCall assumes something might have happened if someone misses calls, but often, people don't answer calls simply because they don't like talking to AI, not because something bad happened. We now have to try calling them a few times and even ask local public officials to visit when they miss CareCall. But honestly, I'm skeptical about putting in so much effort."* P-SiteI, who manages elderly veterans, noted that lapses were often due to forgetfulness, rather than emergencies: *"We send out CareCall calls at a set time every week, so we ask them to be on the lookout. However, elderly individuals often leave their phones somewhere and forget. We keep calling them until they answer, but there's not much we can do if people don't pick up the phone."* To minimize redundant tasks, frontline workers suggested expanding AI's role by further automating follow-ups. P-SiteJ stated: *"For now, we're responsible for reaching out to those who didn't answer CareCall, but I wish it could try again the next day automatically."* P-SiteI similarly wished, *"the system requires minimal or no human intervention. It could try a few more times without prompting us to follow up manually."*

In addition, frontline workers noted that lapses in user engagement with CareCall created additional tasks of handling callbacks. Since the chatbot used frontline workers' office numbers, participants frequently returned calls to their offices after seeing missed calls from CareCall. However, frontline workers frequently found these callbacks overburdening. P-SiteG explained: *"I frequently get callbacks, and people often want to chat while they are at it. No wonder, given we recruited people needing emotional support. But it adds up quickly and becomes overwhelming."* P-SiteB-1 similarly described the difficulty of handling callbacks: *"I get a lot of callbacks from CareCall users. The thing is, it's difficult to keep the call short because they often want to chat and ask about social services, which can easily take 30 minutes per call."*

5.4.4 Creation of a Window to Communicate Different Care Needs. We further found that introducing CareCall unexpectedly served as a window for care recipients to communicate different needs, allowing frontline workers to provide the necessary support. The open-ended nature of LLM-driven chatbots led users to express various healthcare and social service needs when interacting with CareCall. Although CareCall was not targeted at processing such requests, frontline workers valued the ability to identify these needs through call logs and often took action on them.

Frontline workers occasionally discovered individuals' mental health needs through CareCall call logs, which led them to check in on them more frequently. P-SiteA-2 stated: *"We make sure to visit and check in with them if people frequently mention feeling depressed and lonely during CareCall calls. We then connect them to mental health support or job search assistance when necessary."* Working at a mental health center, P-SiteE-1 particularly paid attention to expressions of emotional distress when monitoring CareCall calls and followed up when concerns arose: *"We call when we find something stands out in the call logs. For example, if someone says, 'I'm so depressed. I* *just want to die,' I make sure to reach out. In many cases, just listening to whatever they want to say makes them feel better."*

In addition, in a community health center setting, frontline workers were able to encourage individuals to seek clinical care when they noticed physical health concerns from the call logs. P-SiteC-2 described taking action on an individual's health issue through monitoring CareCall logs: *"Many elderly people take sleeping pills when they frequently wake up at night, but in fact, these issues are often due to urological issues. When we noticed someone repeatedly mentioning their sleep problems, I followed up and encouraged them to visit us to see a urologist. After the visit, thankfully, their sleep issues were resolved."* P-SiteC-1 similarly explained how health concerns expressed during CareCall calls helped them connect to necessary healthcare: *"I noticed someone mentioning severe back pain during the calls. I followed up and encouraged them to visit us so that we could get an X-ray and provide some physical therapy."*

Further, frontline workers identified and addressed social service needs upon monitoring CareCall logs, even though those were beyond the intended scope. P-SiteB-3 explained their desire to connect individuals to social services through CareCall monitoring: *"We know that CareCall is just for check-ins, but we wanted more from the start. Like, connecting people to relevant social services if needs arise, such as job searching or sending home aides."* P-SiteB-2 added that they would note details like, *"someone had surgery recently and asked for financial aid"* when monitoring call logs so local public officers can connect them to resources. P-SiteL described how CareCall turned out useful for natural disaster recovery: *"Last monsoon season was pretty bad in our city, and people mentioned issues like water leaks or floods during CareCall calls. During the season, I paid extra attention to the call logs and reached out to the community centers for help, like fixing their houses. I know the system isn't meant for emergency responses, but it worked out well."*

6 DISCUSSION

Our findings reveal a discrepancy between public agencies' expectations and the realities of deploying CareCall, as it required significant human effort to manage the expanded care, contrary to expectations that it would alleviate the burden of frontline public health monitoring. This mismatch between perceptions and reality of the adoption of LLMs parallels conversations in other spaces, where the technology is often assumed to have greater capabilities than it can deliver on [27, 28, 72]. Looking at our findings from the perspective of articulation work [71], we highlight that decision-makers in public health face unique challenges in conducting articulation work required for AI chatbot adoption, particularly due to the open-ended nature of LLM-driven chatbots. We point to the need to develop guidelines and best practices for decision-makers implementing these emerging technologies in public health contexts. Our findings also surface that CareCall introduced significant maintenance work for frontline workers, primarily due to unmet expectations around user engagement with these chatbots. We argue for the importance of acknowledging and accounting for the maintenance work that these AI chatbots demand in public health monitoring. In addition, we provide implications for public agencies considering the use of AI chatbots for public health monitoring, focusing on the potential of open-ended conversations to identify unmet care needs and the need to assess the impacts of AI adoption on the labor demands of their workforce. For designers and developers aiming to make AI chatbots usable for public health monitoring, we suggest opportunities to piggyback on public infrastructure, incorporate fallback mechanisms to address lapses, and leverage passive sensing to complement chatbots. Finally, we report on the limitations of the study, particularly concerning the transferability of the findings to different countries and domains.

6.1 Considering Decision-Makers' Articulation Work for AI Chatbot Adoption

Through this study, we found that decision-makers in public agencies often expected AI chatbots to reach more people. These expectations were largely realized, but decision-makers often failed to plan adequately for the necessary resources to operate these systems effectively. Overall, these AI chatbots failed to deliver on the expectation of scaling up monitoring reach and frequency without needing to increasing staff. Decision-makers initially perceived AI chatbots like CareCall as highly efficient and requiring minimal human oversight, having unrealistic expectations about their capabilities for public health monitoring. Such perceptions might have been influenced by their personal interactions with open-domain, LLM-driven chatbots (e.g., ChatGPT, Google Gemini) performing naturalistic conversations on diverse topics, as well as broader media and cultural conversations around these technologies. However, in reality, frontline workers and administrators had to take on significant human effort to manage the expanded care with these AI chatbots, such as following up on health concerns and handling lapses. This discrepancy led to failures in planning for the human resources required to operate AI for large-scale health monitoring, adding extra work to already overburdened frontline workers who perform crucial care work with limited resources.

Extending prior CSCW research that sheds light on the articulation work required for introducing new technology into complex healthcare infrastructures [4, 22, 66, 71], our findings highlight that decision-makers may face greater challenges in conducting articulation work for allocating resources for LLM-driven chatbots due to the unrealistic expectations about the capabilities of these technologies. The lack of established guidelines and best practices for these emerging technologies in public health space further introduces greater uncertainty about their capabilities and limitations, as well as the human efforts necessary to operate them. Given these challenges, developing guidelines and best practices for decision-makers in implementing LLM-driven chatbots in public health contexts is a valuable direction for future research. Similar to the toolkit proposed by a recent study [35], such resources could provide valuable insights to support the articulation work of decision-makers and help them develop more realistic expectations of the capabilities of these systems. By aiding them in navigating the complexities of implementing emerging technologies like LLM-driven chatbots, decision-makers could better estimate the human efforts necessary to operate these systems in a more sustainable manner.

6.2 Accounting for Maintenance Work AI Chatbots Impose on Frontline Workers

It can be expected that any change to existing practices in complex public health contexts will introduce new types of labor. However, in this study, we found that the adoption of CareCall introduced not only new labor but also labor that decision-makers in public agencies did not anticipate, such as handling calls that the AI chatbot made but were not responded to. Consistent with prior work [64, 72, 78], our findings reveal that the frontline and administrative workers had to perform substantial maintenance work due to AI chatbots failing to meet some expectations the decision-makers had. The adoption of CareCall for public health monitoring operated under the assumption that missed calls indicate potential health emergencies of users, implying that users would mostly answer the calls when not in emergencies. However, in reality, our frontline workers thought that users frequently lapsed in the use of CareCall because they simply forgot about it or did not want to interact with AI. This suggests a misunderstanding among decision-makers that people would consistently and willingly communicate with AI chatbots in public health monitoring contexts. Research on Personal Informatics has shown that people commonly lapse in the use of health monitoring technology in general, both intentionally and unintentionally [13, 15]. In addition, prior studies have highlighted that people are often hesitant to interact with chatbots, preferring instead to engage with humans behind chatbot-based health interventions [33, 45, 63]. Building on prior work, our study suggests that public health monitoring involving AI chatbots inevitably requires maintenance work to address lapses, whether due to people's patterns of behavior with health monitoring technology in general or their reluctance to interact with AI chatbots.

Previous studies have pointed out that technology could be used as a means to demand additional expectations for frontline workers [76] or normalize increasing their workloads in the name of innovation [72]. Extending prior work, our study suggests that while integrating AI chatbots likely introduces substantial maintenance work to address the limitations of these tools, decision-makers can easily overlook them under the guise of innovation, burdening frontline workers with additional and potentially redundant tasks. Reflecting prior work [3, 53, 78], our findings suggest that an important avenue for improving AI adoption in public health monitoring is to adequately recognize the efforts of workers in maintaining care infrastructure during breakdowns. To avoid overburdening the public health workforce, we recommend that decision-makers pay close attention to the types of tasks AI could add to care infrastructure, who will perform the additional and potentially invisible work, and how such work can be better recognized.

6.3 Implications for Public Agencies

In this section, we highlight implications for public agencies looking to leverage AI chatbots for public health monitoring. Our findings surface that CareCall unexpectedly served as a conduit for communicating various social service and healthcare needs. As an open-domain, LLM-driven chatbot, CareCall supports free-form conversations on serendipitous topics users bring up, allowing users to convey various care needs that traditional task-oriented systems with pre-defined conversation flows might miss [47]. Our study suggests that open-domain, LLM-driven chatbots can play a valuable role in care infrastructures as the safety net for vulnerable populations. We recommend that public agencies **leverage the public's open-ended interactions with LLM-driven chatbots to uncover unmet care needs**. When integrating these chatbots into care infrastructures, it would be essential to establish comprehensive mechanisms to monitor unmet care needs and refer them to relevant social or healthcare services.

However, previous studies highlighted that stakeholders often have unrealistic expectations towards AI systems in the public sector and ascribe more capabilities than they actually can offer [30, 35, 59]. In addition to decision-makers overestimating their capabilities of LLM-driven chatbots, the public may be similarly influenced by personal experience and public discourse around these technologies, potentially leading to disappointment when they cannot receive the care they desire [30]. To maintain realistic expectations of these chatbots, it is crucial for public agencies to transparently communicate system capabilities and limitations with end-users and clarify what public agencies can and cannot offer by monitoring the data collected through these systems.

Further, our findings suggest that adopting AI chatbots necessitates rethinking and reconfiguring the labor involved in public health monitoring. Before introducing technological interventions, frontline public health monitoring primarily involved giving calls or visiting homes to ask routine questions about the health and wellbeing of individuals under their care. In contrast, with CareCall, frontline workers shifted their focus to following up on those who expressed health concerns and addressing lapses. In short, the AI chatbot took over some of the labor expected of frontline workers and introduced different labor in its place. If these AI technologies are widely put into practice for public health monitoring, what we observed in CareCall suggests that the types of labor performed by frontline workers will shift, as well as the kinds of expertise and training needed for these workers to be effective. Our participants frequently stated that they felt that working with CareCall required tasks they had not been trained to do or were well outside their areas of expertise. We do not aim to argue that this shift is positive or negative as a whole. Rather, we urge public agencies to **critically assess the impacts that introducing these AI systems will have on the day-to-day practices of their workforce** as part of deciding whether and how to adopt them.

A core question that decision-makers in public agencies need to consider is the various costs (e.g., financial, morale) associated with re-training frontline workers or hiring new ones to manage the work introduced by AI chatbots. However, given that the day-to-day experiences of multiple stakeholders are impacted by such decisions, we see a need for decision-makers to better acknowledge the new labor placed on existing frontline workers. Prior work has pointed out the clear solution of increasing worker compensation or hiring additional workers [53], though this approach often faces barriers in typically underfunded public agencies. Other approaches include increasing the visibility of this additional work [54, 78], such as through time-tracking, though these can lead to feelings of surveillance. More participatory approaches, involving all stakeholders in the conversation of whether and how to adopt an AI chatbot for public health monitoring, can potentially address some of these concerns and risks [6, 18, 55, 69].

6.4 Implications for Designers and Developers

In this section, we offer implications for designers and developers aiming to make AI chatbots usable for public health monitoring. One significant factor that enabled CareCall to meet stakeholder expectations around expanding reach was it piggybacked on existing public infrastructure, specifically telephone networks. Unlike technologies requiring dedicated hardware, CareCall's chatbot interactions were conducted via phone calls using existing telephone lines. This approach significantly lowered costs, allowing broader reach within public agencies' budget constraints, and reduced the burden on frontline workers and administrators, who would otherwise have been tasked with managing hardware. Research in the HCI community has increasingly underscored the need to consider the scalability of health technologies [51, 79]. When developing chatbots for large-scale health monitoring, building on existing infrastructure [21, 75, 77] or social platforms [14, 24] can enhance scalability, as this approach can lower development and management burdens [14] and facilitate broader engagement in low-resource settings [77]. Consistent with prior work, our study highlights **piggybacking on public infrastructure as a promising strategy to address the scalability challenges** of public health monitoring chatbots. When deciding whether and what public infrastructure to piggyback on, designers need to carefully consider its impact on end-user interactions and broader stakeholder workflows.

In addition, our study points to opportunities for developers to **incorporate fallback mechanisms to address lapses in user engagement** with chatbot-based public health monitoring. While lapses are to be expected [13, 15], they are often respected in other contexts. However, in critical health contexts where public health monitoring is often deployed, such as for the prevention of lonely deaths, such lapses could indicate serious health emergencies or even death, and it may not be beneficial to outright ignore them. Incorporating fallback mechanisms can help make chatbot-based public health monitoring more resilient to lapses in user engagement in the long term. Prior work pointed to the opportunities for using secondary sources that generate data as a byproduct of the daily digital lives of individuals—such as social media posts and app usage [11, 13, 57, 73]—or in-home environment monitoring—such as water usage or electricity consumption [21, 42, 75]. While our findings revealed that public health officials often view such passive sensing approaches as too error-prone to serve as the primary method of monitoring personal health, they could be effective as fallback mechanisms for addressing lapses in user engagement. These approaches, as secondary sources of producing data, do not require additional effort from individuals, reducing frontline workers' burden of handling lapses in chatbot-based public health monitoring. One important factor to consider when leveraging passive sensing approaches as fallback mechanisms for public health monitoring is whether dedicated hardware devices are required because it likely introduces additional labor to maintain them and limits public reach due to the cost. We suggest that developers carefully evaluate the opportunity to leverage existing public infrastructure as a fallback mechanism to chatbot-based public health monitoring, while also considering whether approaches that do not build on existing infrastructure could provide additional value.

Beyond addressing lapses in user engagement, we further see opportunities for **passive sensing to complement chatbots to align with various public health monitoring needs**. Public officials highlighted how individuals often had skepticism or concerns around the use of conversational AI for monitoring, and providing an alternative technical approach helped them monitor this group while aligning with their preferences. Further, officials highlighted that some chronic conditions common in their population, such as hearing loss, were not amenable to voice-based chatbot check-ins, and having an alternative was beneficial. Beyond serving as an alternative, there are likely opportunities for passive sensing approaches to deepen understanding of the public's daily experiences. For example, understanding energy consumption patterns via in-home sensors could triangulate self-reported behaviors via chatbots, creating a better picture of how an individual living alone is doing. However, our findings suggest that public agencies often adopt new technological interventions sequentially given resource constraints, evolving policy priorities, and emerging technology trends, so care must be taken when designing integrated approaches in ways that can be readily adopted.

6.5 Limitations and Future Work

Our goal for this study was to understand the expectations and realities faced by public agencies deploying AI chatbots for public health monitoring through the case of CareCall, which led us to focus on the perspectives of public agency workers who were involved in its adoption and rollout in South Korea. We believe that many of the circumstances that our interviewees described—such as the lack of resources in public agencies and the demanding working conditions of frontline workers—are not unique to the South Korean context or the specific deployment of CareCall, suggesting that similar expectations and challenges could arise when AI chatbots are rolled out for public health monitoring in other countries. However, we acknowledge that various country-specific factors—such as regulatory requirements, cultural norms, technological infrastructure, public trust in AI systems, and the overall maturity of digital health initiatives—can significantly influence how public agencies approach AI adoption and deployment for public health monitoring. As such, it is crucial to consider these contextual differences when applying our findings to other regions.

Our study focused on the context of CareCall, a system designed to monitor the health and wellbeing of socially isolated individuals, primarily low-SES middle-aged and older adults living alone. Using AI chatbots in public health monitoring for different and broader populations—such as crisis management or chronic disease monitoring—likely involves different interpersonal and infrastructural dynamics. For instance, crisis management often involves real-time monitoring and decision-making, which may involve different labor demands compared to the relatively stable, routine monitoring of individuals' wellbeing. Further, most CareCall deployments in our study context were pilot projects implemented on relatively small scales. When AI chatbots are deployed on a larger scale (e.g., state-wide or nationwide), public agencies may engage in more robust resource planning and role assignment than what was seen in our study. Future research should explore how these factors play out in different public health domains and at various scales of deployment to better understand the perspectives and practices around AI adoption and rollout.

Finally, we recognize that participants were describing their expectations for CareCall retrospectively, and the descriptions they provided were likely influenced by their actual experience with the system. Nonetheless, we expect that participant descriptions were fairly reliable, as they largely lined up with typical expectations that AI technology can help reduce workload and expand scale. Further work would benefit from investigating how public health workers' perceptions of AI chatbots change as they more deeply understand the technology's capabilities, such as participatory methods with longitudinal engagement.

7 CONCLUSION

Through interviews with 21 public agency workers involved in the adoption and deployment of CareCall across decision-making, administration, and frontline monitoring roles, we found that public agencies' expectations for AI chatbots to expand reach were largely met, but frontline workers often experienced an increased burden due to insufficient resources and new labor demands, such as handling lapses in user engagement. Our findings suggest that the open-ended nature of LLM-driven chatbots and the lack of established guidelines around these emerging technologies introduce unique challenges for decision-makers when conducting the articulation work required for AI chatbot implementation. We also highlight the importance of recognizing the maintenance work that AI chatbots impose on frontline workers, especially considering end-user lapses in using these systems. For public agencies, we suggest leveraging open-ended conversations of LLM-driven chatbots to identify unmet care needs and critically assess the impacts of AI implementation on the labor demands of their workforce. For developers, we suggest piggybacking on public infrastructure, incorporating fallback mechanisms to better address lapses in user engagement with AI chatbots, and leveraging passive sensing to complement AI chatbots for public health monitoring.

ACKNOWLEDGMENTS

We thank our participants for providing valuable insights. This work was partially supported by the National Science Foundation under award IIS-2237389 and through a research internship at NAVER AI Lab of NAVER Cloud. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding organizations.

REFERENCES

[1] Sanghwan Bae, Donghyun Kwak, Soyoung Kang, Min Young Lee, Sungdong Kim, Yuin Jeong, Hyeri Kim, Sang-Woo Lee, Woomyoung Park, and Nako Sung. 2022. Keep Me Updated! Memory Management in Long-term Conversations. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang (Eds.). Association for Computational Linguistics, Abu Dhabi, United Arab Emirates, 3769–3787. https://doi.org/10.18653/v1/2022.findings-emnlp.276

[2] Sanghwan Bae, Donghyun Kwak, Sungdong Kim, Donghoon Ham, Soyoung Kang, Sang-Woo Lee, and Woomyoung Park. 2022. Building a Role Specified Open-Domain Dialogue System Leveraging Large-Scale Language Models. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Seattle, United States, 2128–2150. https://doi.org/10.18653/v1/2022.naacl-main.155

[3] Vince Bartle, Janice Lyu, Freesoul El Shabazz-Thompson, Yunmin Oh, Angela Anqi Chen, Yu-Jan Chang, Kenneth Holstein, and Nicola Dell. 2022. "A Second Voice": Investigating Opportunities and Challenges for Interactive Voice Assistants to Support Home Health Aides. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 45, 17 pages. https://doi.org/10.1145/3491102.3517683

[4] Marc Berg. 1999. Patient care information systems and health care work: a sociotechnical approach. *International Journal of Medical Informatics* 55, 2 (1999), 87–101. https://doi.org/10.1016/S1386-5056(99)00011-8

[5] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health* 11, 4 (2019), 589–597.

[6] Anna Brown, Alexandra Chouldechova, Emily Putnam-Hornstein, Andrew Tobin, and Rhema Vaithianathan. 2019. Toward Algorithmic Accountability in Public Services: A Qualitative Study of Affected Community Perspectives on Algorithmic Decision-making in Child Welfare Services. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300271

[7] Hye-jin Byun. 2022. NAVER launches AI call service aimed at seniors - The Korea Herald. Retrieved Sep 14, 2023 from https://www.koreaherald.com/view.php?ud=20220530000643

[8] Yuchen Chen, Yuling Sun, and Silvia Lindtner. 2023. Maintainers of Stability: The Labor of China's Data-Driven Governance and Dynamic Zero-COVID. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 159, 16 pages. https://doi.org/10.1145/3544548.3581299

[9] Hao-Fei Cheng, Logan Stapleton, Ruiqi Wang, Paige Bullock, Alexandra Chouldechova, Zhiwei Steven Steven Wu, and Haiyi Zhu. 2021. Soliciting Stakeholders' Fairness Notions in Child Maltreatment Predictive Systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 390, 17 pages. https://doi.org/10.1145/3411764.3445308

[10] Marion Crain, Winifred Poster, and Miriam Cherry. 2016. *Invisible labor: Hidden work in the contemporary world*. Univ of California Press.

[11] Munmun De Choudhury, Mrinal Kumar, and Ingmar Weber. 2017. Computational Approaches Toward Integrating Quantified Self Sensing and Social Media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 1334–1349. https://doi.org/10.1145/2998181.2998219

[12] Brian DeRenzi, Nicola Dell, Jeremy Wacksman, Scott Lee, and Neal Lesh. 2017. Supporting Community Health Workers in India through Voice- and Web-Based Feedback. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2770–2781. https://doi.org/10.1145/3025453.3025514

[13] Daniel A Epstein, Parisa Eslambolchilar, Judy Kay, Jochen Meyer, and Sean A Munson. 2021. *Opportunities and challenges for long-term tracking*. Springer International Publishing.

[14] Daniel A Epstein, Fannie Liu, Andrés Monroy-Hernández, and Dennis Wang. 2022. Revisiting piggyback prototyping: Examining benefits and tradeoffs in extending existing social computing systems. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–28. https://doi.org/10.1145/3555557

[15] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 731–742. https://doi.org/10.1145/2750858.2804250

[16] Thomas Erickson, Ming Li, Younghun Kim, Ajay Deshpande, Sambit Sahu, Tian Chao, Piyawadee Sukaviriya, and Milind Naphade. 2013. The dubuque electricity portal: evaluation of a city-scale residential electricity consumption feedback system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1203–1212. https://doi.org/10.1145/2470654.2466155

[17] Biyi Fang, Qiumin Xu, Taiwoo Park, and Mi Zhang. 2016. AirSense: an intelligent home-based sensing system for indoor air quality analytics. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Heidelberg, Germany) *(UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 109–119. https://doi.org/10.1145/2971648.2971720

[18] Asbjørn Flügge, Thomas Hildebrandt, and Naja Holten Møller. 2021. Street-level algorithms and AI in bureaucratic decision-making: A caseworker perspective. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1, Article 40 (2021), 23 pages. https://doi.org/10.1145/3449114

[19] Clark C. Freifeld, Rumi Chunara, Sumiko R. Mekaru, Emily H. Chan, Taha Kass-Hout, Anahi Ayala Iacucci, and John S. Brownstein. 2010. Participatory Epidemiology: Use of Mobile Phones for Community-Based Health Reporting. *PLoS Medicine* 7, 12 (Dec. 2010), e1000376. https://doi.org/10.1371/journal.pmed.

1000376

[20] Jon Froehlich, Leah Findlater, Marilyn Ostergren, Solai Ramanathan, Josh Peterson, Inness Wragg, Eric Larson, Fabia Fu, Mazhengmin Bai, Shwetak Patel, and James A. Landay. 2012. The design and evaluation of prototype eco-feedback displays for fixture-level water usage data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2367–2376. https://doi.org/10.1145/2207676.2208397

[21] Jon E. Froehlich, Eric Larson, Tim Campbell, Conor Haggerty, James Fogarty, and Shwetak N. Patel. 2009. HydroSense: infrastructure-mediated single-point sensing of whole-home water activity. In *Proceedings of the 11th International Conference on Ubiquitous Computing* (Orlando, Florida, USA) *(UbiComp '09)*. Association for Computing Machinery, New York, NY, USA, 235–244. https://doi.org/10.1145/1620545.1620581

[22] Elihu M Gerson and Susan Leigh Star. 1986. Analyzing due process in the workplace. *ACM Transactions on Information Systems (TOIS)* 4, 3 (1986), 257–270. https://doi.org/10.1145/214427.214431

[23] Mitchell Gordon, Tim Althoff, and Jure Leskovec. 2019. Goal-setting And Achievement In Activity Tracking Apps: A Case Study Of MyFitnessPal. In *The World Wide Web Conference* (San Francisco, CA, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 571–582. https://doi.org/10.1145/3308558.3313432

[24] Catherine Grevet and Eric Gilbert. 2015. Piggyback Prototyping: Using Existing, Large-Scale Social Computing Systems to Prototype New Ones. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 4047–4056. https://doi.org/10.1145/2702123.2702395

[25] Yen-Chia Hsu, Paul Dille, Jennifer Cross, Beatrice Dias, Randy Sargent, and Illah Nourbakhsh. 2017. Community-Empowered Air Quality Monitoring System. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 1607–1619. https://doi.org/10.1145/3025453.3025853

[26] Azra Ismail and Neha Kumar. 2018. Engaging Solidarity in Data Collection Practices for Community Health. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 76 (Nov. 2018), 24 pages. https://doi.org/10.1145/3274345

[27] Azra Ismail and Neha Kumar. 2021. AI in Global Health: The View from the Front Lines. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–21. https://doi.org/10.1145/3411764.3445130

[28] Azra Ismail, Divy Thakkar, Neha Madhiwalla, and Neha Kumar. 2023. Public Health Calls for/with AI: An Ethnographic Perspective. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 354 (oct 2023), 26 pages. https://doi.org/10.1145/3610203

[29] Azra Ismail, Deepika Yadav, Meghna Gupta, Kirti Dabas, Pushpendra Singh, and Neha Kumar. 2022. Imagining Caring Futures for Frontline Health Work. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (Nov. 2022), 1–30. https://doi.org/10.1145/3555581

[30] Eunkyung Jo, Daniel A. Epstein, Hyunhoon Jung, and Young-Ho Kim. 2023. Understanding the Benefits and Challenges of Deploying Conversational AI Leveraging Large Language Models for Public Health Intervention. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 18, 16 pages. https://doi.org/10.1145/3544548.3581503

[31] Eunkyung Jo, Yuin Jeong, Sohyun Park, Daniel A. Epstein, and Young-Ho Kim. 2024. Understanding the Impact of Long-Term Memory on Self-Disclosure with Large Language Model-Driven Chatbots for Public Health Intervention. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 440, 21 pages. https://doi.org/10.1145/3613904.3642420

[32] Naveena Karusala, Sohini Upadhyay, Rajesh Veeraraghavan, and Krzysztof Z. Gajos. 2024. Understanding Contestability on the Margins: Implications for the Design of Algorithmic Decision-making in Public Services. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 478, 16 pages. https://doi.org/10.1145/3613904.3641898

[33] Naveena Karusala, Shirley Yan, Nupoor Rajkumar, Victoria G, and Richard Anderson. 2023. Speculating with Care: Worker-centered Perspectives on Scale in a Chat-based Health Information Service. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2, Article 361 (Oct. 2023), 26 pages. https://doi.org/10.1145/3610210

[34] Anna Kawakami, Amanda Coston, Hoda Heidari, Kenneth Holstein, and Haiyi Zhu. 2024. Studying Up Public Sector AI: How Networks of Power Relations Shape Agency Decisions Around AI Design and Use. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2, Article 450 (Nov. 2024), 24 pages. https://doi.org/10.1145/3686989

[35] Anna Kawakami, Amanda Coston, Haiyi Zhu, Hoda Heidari, and Kenneth Holstein. 2024. The Situate AI Guidebook: Co-Designing a Toolkit to Support Multi-Stakeholder, Early-stage Deliberations Around Public Sector AI Proposals. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 749, 22 pages. https://doi.org/10.1145/3613904.3642849

[36] Anna Kawakami, Venkatesh Sivaraman, Hao-Fei Cheng, Logan Stapleton, Yanghuidi Cheng, Diana Qing, Adam Perer, Zhiwei Steven Wu, Haiyi Zhu, and Kenneth Holstein. 2022. Improving Human-AI Partnerships in Child Welfare: Understanding Worker Practices, Challenges, and Desires for Algorithmic Decision Support. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 52, 18 pages. https://doi.org/10.1145/3491102.3517439

[37] Anna Kawakami, Venkatesh Sivaraman, Logan Stapleton, Hao-Fei Cheng, Adam Perer, Zhiwei Steven Wu, Haiyi Zhu, and Kenneth Holstein. 2022. "Why Do I Care What's Similar?" Probing Challenges in AI-Assisted Child Welfare Decision-Making through Worker-AI Interface Design Concepts. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 454–470. https://doi.org/10.1145/3532106.3533556

[38] Elizabeth Kaziunas, Michael S. Klinkman, and Mark S. Ackerman. 2019. Precarious Interventions: Designing for Ecologies of Care. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW, Article 113 (Nov. 2019), 27 pages. https://doi.org/10.1145/3359215

[39] Jackson A. Killian, Bryan Wilder, Amit Sharma, Vinod Choudhary, Bistra Dilkina, and Milind Tambe. 2019. Learning to Prescribe Interventions for Tuberculosis Patients Using Digital Adherence Data. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Anchorage, AK, USA) *(KDD '19)*. Association for Computing Machinery, New York, NY, USA, 2430–2438. https://doi.org/10.1145/3292500.3330777

[40] Boseop Kim, HyoungSeok Kim, Sang-Woo Lee, Gichang Lee, Donghyun Kwak, Jeon Dong Hyeon, Sunghyun Park, Sungju Kim, Seonhoon Kim, Dongpil Seo, Heungsub Lee, Minyoung Jeong, Sungjae Lee, Minsub Kim, Suk Hyun Ko, Seokhun Kim, Taeyong Park, Jinuk Kim, Soyoung Kang, Na-Hyeon Ryu, Kang Min Yoo, Minsuk Chang, Soobin Suh, Sookyo In, Jinseong Park, Kyungduk Kim, Hiun Kim, Jisu Jeong, Yong Goo Yeo, Donghoon Ham, Dongju Park, Min Young Lee, Jaewook Kang, Inho Kang, Jung-Woo Ha, Woomyoung Park, and Nako Sung. 2021. What Changes Can Large-scale Language Models Bring? Intensive Study on HyperCLOVA: Billions-scale Korean Generative Pretrained Transformers. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Online and Punta Cana, Dominican Republic, 3405–3424. https://doi.org/10.18653/v1/2021.emnlp-main.274

[41] Junhan Kim, Jana Muhic, Lionel Peter Robert, and Sun Young Park. 2022. Designing Chatbots with Black Americans with Chronic Conditions: Overcoming Challenges against COVID-19. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 439, 17 pages. https://doi.org/10.1145/3491102.3502116

[42] Wilhelm Kleiminger, Christian Beckel, and Silvia Santini. 2015. Household occupancy monitoring using electricity meters. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 975–986. https://doi.org/10.1145/2750858.2807538

[43] Korea Law Translation Center. 2020. Act on the Prevention and Management of Lonely Deaths. https://elaw.klri.re.kr/eng_mobile/viewer.do?hseq=55028&type=part&key=38

[44] Tzu-Sheng Kuo, Hong Shen, Jisoo Geum, Nev Jones, Jason I. Hong, Haiyi Zhu, and Kenneth Holstein. 2023. Understanding Frontline Workers' and Unhoused Individuals' Perspectives on AI Used in Homeless Services. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 860, 17 pages. https://doi.org/10.1145/3544548.3580882

[45] Amanda Lazar, Hilaire J. Thompson, Shih Yin Lin, and George Demiris. 2018. Negotiating relation work with telehealth home care companionship technologies that support aging in place. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW, Article 103 (Nov. 2018), 19 pages. https://doi.org/10.1145/3274372

[46] Charlotte P. Lee, Paul Dourish, and Gloria Mark. 2006. The human infrastructure of cyberinfrastructure. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work* (Banff, Alberta, Canada) *(CSCW '06)*. Association for Computing Machinery, New York, NY, USA, 483–492. https://doi.org/10.1145/1180875.1180950

[47] Minha Lee, Sander Ackermans, Nena van As, Hanwen Chang, Enzo Lucas, and Wijnand IJsselsteijn. 2019. Caring for Vincent: A Chatbot for Self-Compassion. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300932

[48] Yeonjin Lee and Doo-Sub Kim. 2020. *Internal Migration in South Korea.* Springer International Publishing, Cham, 93–111. https://doi.org/10.1007/978-3-030-44010-7_6

[49] Xi Lu, Eunkyung Jo, Seora Park, Hwajung Hong, Yunan Chen, and Daniel A Epstein. 2022. Understanding cultural influence on perspectives around contact tracing strategies. *Proceedings of the ACM on Human-Computer Interaction* 6,

CSCW2 (2022), 1–26. https://doi.org/10.1145/3555569

[50] Xi Lu, Tera L. Reynolds, Eunkyung Jo, Hwajung Hong, Xinru Page, Yunan Chen, and Daniel A. Epstein. 2021. Comparing Perspectives Around Human and Technology Support for Contact Tracing. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 200, 15 pages. https://doi.org/10.1145/3411764.3445669

[51] Aaron Lyon, Sean A. Munson, Madhu Reddy, Stephen M. Schueller, Elena Agapie, Svetlana Yarosh, Alex Dopp, Ulrica von Thiele Schwarz, Gavin Doherty, Andrea K Graham, Kaylee Payne Kruzan, and Rachel Kornfield. 2023. Bridging HCI and Implementation Science for Innovation Adoption and Public Health Impact. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 336, 7 pages. https://doi.org/10.1145/3544549.3574132

[52] Aditya Mate, Lovish Madaan, Aparna Taneja, Neha Madhiwalla, Shresth Verma, Gargi Singh, Aparna Hegde, Pradeep Varakantham, and Milind Tambe. 2022. Field Study in Deploying Restless Multi-Armed Bandits: Assisting Non-profits in Improving Maternal and Child Health. *Proceedings of the AAAI Conference on Artificial Intelligence* 36, 11, 12017–12025. https://doi.org/10.1609/aaai.v36i11.21460

[53] Joy Ming, Srujana Kamath, Elizabeth Kuo, Madeline Sterling, Nicola Dell, and Aditya Vashistha. 2022. Invisible Work in Two Frontline Health Contexts. In *Proceedings of the 5th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies* (Seattle, WA, USA) *(COMPASS '22)*. Association for Computing Machinery, New York, NY, USA, 139–151. https://doi.org/10.1145/3530190.3534814

[54] Joy Ming, Elizabeth Kuo, Katie Go, Emily Tseng, John Kallas, Aditya Vashistha, Madeline Sterling, and Nicola Dell. 2023. "I Go Beyond and Beyond" Examining the Invisible Work of Home Health Aides. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1, Article 59 (April 2023), 21 pages. https://doi.org/10.1145/3579492

[55] Naja Holten Møller, Irina Shklovski, and Thomas T. Hildebrandt. 2020. Shifting Concepts of Value: Designing Algorithmic Decision-Support Systems for Public Services. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (Tallinn, Estonia) *(NordiCHI '20)*. Association for Computing Machinery, New York, NY, USA, Article 70, 12 pages. https://doi.org/10.1145/3419249.3420149

[56] Jimmy Moore, Pascal Goffin, Miriah Meyer, Philip Lundrigan, Neal Patwari, Katherine Sward, and Jason Wiese. 2018. Managing In-home Environments through Sensing, Annotating, and Visualizing Air Quality Data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3, Article 128 (Sept. 2018), 28 pages. https://doi.org/10.1145/3264938

[57] Elizabeth L. Murnane, Tara G. Walker, Beck Tench, Stephen Voida, and Jaime Snyder. 2018. Personal Informatics in Interpersonal Contexts: Towards the Design of Technology that Supports the Social Ecologies of Long-Term Mental Health Management. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW, Article 127 (Nov. 2018), 27 pages. https://doi.org/10.1145/3274396

[58] Fabian Okeke, Emily Tseng, Benedetta Piantella, Mikaela Brown, Harveen Kaur, Madeline R. Sterling, and Nicola Dell. 2019. Technology, home health care, and heart failure: a qualitative analysis with multiple stakeholders. In *Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies* (Accra, Ghana) *(COMPASS '19)*. Association for Computing Machinery, New York, NY, USA, 122–133. https://doi.org/10.1145/3314344.3332487

[59] Chinasa T. Okolo, Srujana Kamath, Nicola Dell, and Aditya Vashistha. 2021. "It cannot do all of my work": Community Health Worker Perceptions of AI-Enabled Mobile Health Applications in Rural India. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 701, 20 pages. https://doi.org/10.1145/3411764.3445420

[60] Joyojeet Pal, Anjuli Dasika, Ahmad Hasan, Jackie Wolf, Nick Reid, Vaishnav Kameswaran, Purva Yardi, Allyson Mackay, Abram Wagner, Bhramar Mukherjee, Sucheta Joshi, Sujay Santra, and Priyamvada Pandey. 2017. Changing data practices for community health workers: Introducing digital data collection in West Bengal, India. In *Proceedings of the Ninth International Conference on Information and Communication Technologies and Development* (Lahore, Pakistan) *(ICTD '17)*. Association for Computing Machinery, New York, NY, USA, Article 17, 12 pages. https://doi.org/10.1145/3136560.3136582

[61] Shwetak N. Patel, Sidhant Gupta, and Matthew S. Reynolds. 2010. The design and evaluation of an end-user-deployable, whole house, contactless power consumption sensor. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 2471–2480. https://doi.org/10.1145/1753326.1753700

[62] Sachin R. Pendse, Faisal M. Lalani, Munmun De Choudhury, Amit Sharma, and Neha Kumar. 2020. "Like Shock Absorbers": Understanding the Human Infrastructures of Technology-Mediated Mental Health Support. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376465

[63] Trevor Perrier, Nicola Dell, Brian DeRenzi, Richard Anderson, John Kinuthia, Jennifer Unger, and Grace John-Stewart. 2015. Engaging Pregnant Women in Kenya with a Hybrid Computer-Human SMS Communication System. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 1429–1438. https://doi.org/10.1145/2702123.2702124

[64] Nithya Sambasivan and Thomas Smyth. 2010. The human infrastructure of ICTD. In *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development* (London, United Kingdom) *(ICTD '10)*. Association for Computing Machinery, New York, NY, USA, Article 40, 9 pages. https://doi.org/10.1145/2369220.2369258

[65] Devansh Saxena, Karla Badillo-Urquiola, Pamela J. Wisniewski, and Shion Guha. 2021. A Framework of High-Stakes Algorithmic Decision-Making for the Public Sector Developed through a Case Study of Child-Welfare. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–41. https://doi.org/10.1145/3476089

[66] Kjeld Schmidt and Liam Bannon. 1992. Taking CSCW seriously: Supporting articulation work. Vol. 1. Springer, 7–40. https://doi.org/10.1007/BF00752449

[67] Ali Shameli, Tim Althoff, Amin Saberi, and Jure Leskovec. 2017. How Gamification Affects Physical Activity: Large-scale Analysis of Walking Challenges in a Mobile Application. In *Proceedings of the 26th International Conference on World Wide Web Companion* (Perth, Australia) *(WWW '17 Companion)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 455–463. https://doi.org/10.1145/3041021.3054172

[68] Mark S. Smolinski, Adam W. Crawley, Kristin Baltrusaitis, Rumi Chunara, Jennifer M. Olsen, Oktawia Wójcik, Mauricio Santillana, Andre Nguyen, and John S. Brownstein. 2015. Flu Near You: Crowdsourced Symptom Reporting Spanning 2 Influenza Seasons. *American Journal of Public Health* 105, 10 (Oct. 2015), 2124–2130. https://doi.org/10.2105/AJPH.2015.302696

[69] Logan Stapleton, Min Hun Lee, Diana Qing, Marya Wright, Alexandra Chouldechova, Ken Holstein, Zhiwei Steven Wu, and Haiyi Zhu. 2022. Imagining new futures beyond predictive systems in child welfare: A qualitative study with impacted stakeholders. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1162–1177. https://doi.org/10.1145/3531146.3533177

[70] Susan Star and Karen Ruhleder. 1996. Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. *Information Systems Research* 7 (03 1996), 111–134. https://doi.org/10.1287/isre.7.1.111

[71] Anselm Strauss. 1988. The Articulation of Project Work: An Organizational Process. *The Sociological Quarterly* 29, 2 (June 1988), 163–178. https://doi.org/10.1111/j.1533-8525.1988.tb01249.x

[72] Yuling Sun, Xiaojuan Ma, Silvia Lindtner, and Liang He. 2023. Data Work of Frontline Care Workers: Practices, Problems, and Opportunities in the Context of Data-Driven Long-Term Care. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (April 2023), 1–28. https://doi.org/10.1145/3579475

[73] Vedant Das Swain, Jingjing Ye, Siva Karthik Ramesh, Abhirup Mondal, Gregory D Abowd, Munmun De Choudhury, et al. 2024. Leveraging Social Media to Predict COVID-19–Induced Disruptions to Mental Well-Being Among University Students: Modeling Study. *JMIR Formative Research* 8, 1 (2024), e52316.

[74] Stephen B. Thacker and Ruth L. Berkelman. 1988. Public health surveillance in the United States. *Epidemiologic Reviews* 10, 1 (1988), 164–190. https://doi.org/10.1093/oxfordjournals.epirev.a036021

[75] Edison Thomaz, Vinay Bettadapura, Gabriel Reyes, Megha Sandesh, Grant Schindler, Thomas Plötz, Gregory D. Abowd, and Irfan Essa. 2012. Recognizing water-based activities in the home through infrastructure-mediated sensing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (Pittsburgh, Pennsylvania) *(UbiComp '12)*. Association for Computing Machinery, New York, NY, USA, 85–94. https://doi.org/10.1145/2370216.2370230

[76] Emily Tseng, Fabian Okeke, Madeline Sterling, and Nicola Dell. 2020. "We can learn. Why not?": Designing Technologies to Engender Equity for Home Health Aides. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376633

[77] Aditya Vashistha, Edward Cutrell, Gaetano Borriello, and William Thies. 2015. Sangeet Swara: A Community-Moderated Voice Forum in Rural India. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 417–426. https://doi.org/10.1145/2702123.2702191

[78] Nervo Verdezoto, Naveen Bagalkot, Syeda Zainab Akbar, Swati Sharma, Nicola Mackintosh, Deirdre Harrington, and Paula Griffiths. 2021. The Invisible Work of Maintenance in Community Health: Challenges and Opportunities for Digital Health to Support Frontline Health Workers in Karnataka, South India. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (April 2021), 1–31. https://doi.org/10.1145/3449165

[79] Alex Waddell, Joshua Paolo Seguin, Ling Wu, Peta Stragalinos, Joe Wherton, Jessica L Watterson, Christopher Owen Prawira, Patrick Olivier, Victoria Manning, Dan Lubman, and Jasmin Grigg. 2024. Leveraging Implementation Science in

Human-Centred Design for Digital Health. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 240, 17 pages. https://doi.org/10.1145/3613904.3642161

[80] CNN World. 2022. The lonely deaths: South Korea's government scrambles to tackle the rise of 'godoksa'. Retrieved Sep 14, 2023 from https://www.cnn.com/2022/12/18/asia/south-korea-godoksa-lonely-death-intl-hnk-dst/index.htm

[81] Amulya Yadav, Bryan Wilder, Eric Rice, Robin Petering, Jaih Craddock, Amanda Yoshioka-Maxwell, Mary Hemler, Laura Onasch-Vera, Milind Tambe, and Darlene Woo. 2018. Bridging the gap between theory and practice in influence maximization: Raising awareness about HIV among homeless youth.. In *International Joint Conferences on Artificial Inteligence*. 5399–5403. https://doi.org/10.24963/ijcai.2018/761

[82] Sailin Zhong, Denis Lalanne, and Hamed Alavi. 2021. The Complexity of Indoor Air Quality Forecasting and the Simplicity of Interacting with It – A Case Study of 1007 Office Meetings. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 47, 19 pages. https://doi.org/10.1145/3411764.3445524