Data Work of Frontline Care Workers: Practices, Problems, and Opportunities in the Context of Data-Driven Long-Term Care

YULING SUN, East China Normal University, China
XIAOJUAN MA, Hong Kong University of Science and Technology, Hong Kong SAR
SILVIA LINDTNER, University of Michigan, USA
LIANG HE, East China Normal University, China

Using data and data technologies to support healthcare has drawn significant attention recently. While CSCW and HCI have largely celebrated the tremendous promise of 'data-driven healthcare' in reforming the healthcare sector, this paper reveals 'labor-driven reality' of this promised data-driven future. Drawing from a qualitative study in a real-world data-driven long-term care (LTC) facility in China, we demonstrate how data-driven technologies work in practice, and especially how frontline workers, as the crux of this data-driven configuration, conduct a tremendous amount of "data work" to make data-drivenness work. This data work, we argue, goes beyond the "clerical work" and functions as a labor of maintenance, articulation, and repair, that both guarantees data technologies' functionalities and acts as an interface between stakeholders. We conclude by discussing the practices, problems and opportunities of this data work in a boarder socio-cultural context.

CCS Concepts: • **Human-centered computing** → Collaborative and social computing.

Additional Key Words and Phrases: Data, data work, data-driven healthcare, long-term care, frontline care worker, invisible labor, data-driven healthcare, data-driven society

ACM Reference format:
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. *Proc. ACM Hum.-Comput. Interact,* 7, CSCW1, Article 42 (April 2023), 28 pages, https://doi.org/10.1145/ 3579475

1 INTRODUCTION

With the arrival of aging boom on a worldwide scale and the spread of chronic diseases [58], Long-Term Care (LTC) – a variety of care services designed to meet both health and personal needs of people who can't care for themselves for long periods [56] – is rapidly becoming an urgent topic in today's society [37, 67]. Recently, considerable academic and industrial efforts

---

Author's addresses: Yuling Sun, ylsun@cs.ecnu.edu.cn, Shanghai Institute of AI for Education and School of Computer Science and Technology, East China Normal University, Shanghai, China; Xiaojuan Ma, mxj@cse.ust.hk, Hong Kong University of Science and Technology, Hong Kong SAR; Silvia Lindtner, lindtner@umich.edu, University of Michigan, Michigan, USA; Liang He (corresponding author), lhe@cs.ecnu.edu.cn, School of Computer Science and Technology, East China Normal University, Shanghai, China.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
2573-0142/2023/04 - ART42 $15.00
© Copyright is held by the owner/author(s). Publication rights licensed to ACM.
https://doi.org/10.1145/3579475

have been paid to design and use various data technologies to enable data-driven LTC reform and address unprecedented-related LTC stress and challenges.

While data-driven LTC has been largely celebrated due to its tremendous promise in, for instance, improving care quality, controlling soaring costs, and supporting better management [55, 76], increasing recent literature has reported that what makes 'data-drivenness' become true in healthcare is not a specific data system, technology, or platform, but a broad chain of socio-technological practices and processes between human and data [6, 40, 63], such as creating, collecting, recording, managing, curating, analyzing, interpreting, contextualizing, communicating data etc. These practices, defined as "data work" [6], root in particular places and times and often require considerable human efforts. Researchers emphasize that it is crucial to better understand data work in various healthcare settings for improving the practical outcome of data-driven healthcare [6, 47].

In LTC settings, due to the extremely long duration, and complex and fragmented nature [10, 56], this data work is more challenging. Often, people who need LTC services are frail, older, and have both physical and cognitive impairments, or disabilities with functional limitations. Their daily lives rely almost entirely on the services of care-providers, including not only medical and health care, but also personal, social, and emotional care [16, 56]. The needed data thus range far beyond physiological data in general electronic health record system (EHR) to much socio-psychological, emotional and ecological data [39, 77], most of which are personalized, situated, unstructured, and hard to collect and format [60]. Further, the collaborations in LTC are extremely complex, encompassing numerous synchronous and asynchronous workflows [72] among numerous occupations and workers, e.g. formal and informal caregivers, nurses, doctors, therapists, social workers, etc. These unique characteristics result in tremendous difficulties in producing, collecting, formatting, analyzing, and using data in LTC setting. In CSCW and HCI, although there has been surging interest in data-driven LTC, issues related to the practical sociotechnical data practices in 'data-driven LTC' are largely underexplored.

To fill this gap, we conducted a qualitative study with 19 frontline care workers in data-driven LTC facilities in China. The term 'frontline care worker' in this paper refers to permanent care staff who are paid to provide day-to-day care for care-receivers. We chose frontline care workers as our study participants because they are the actual performers of most day-to-day care work in LTC setting and the primary users and handlers of most data technologies. Drawing on the notion of "data work [6]", our goal, here, then is honing in on the data practices, impacts and challenges of frontline care workers in LTC. We specifically ask the following three research questions (RQs):

RQ1: *How integrated is data-driven healthcare in the LTC process? How does it work in practice?*

RQ2: *What and how is data work performed by frontline care workers in the contexts of data-driven LTC? Why and for whom?*

RQ3: *How does this data work introduced by 'data-drivenness' impact frontline care workers' practical care-taking work?*

Our findings suggest frontline care workers, who are generally considered as "low-skilled" workers [9, 10, 18], are actually the crux of data-driven LTC configuration. They (have to) provide a tremendous amount of complex and effortful data work, including constructing and maintaining dataset, articulating stakeholders, repairing breakdowns, etc., to make data-driven LTC become true. This data work is often invisible, unacknowledged and unpaid, which causes the increased unfairness. Further, it also triggers some forms of data cascades [62], and causes negative and downstream effects to data-driven healthcare.

Our study contributes to CSCW communities in three specific ways. First, we contribute to ongoing studies examining data work in "data-drivenness" with particular focus on data work of frontline care workers with lower social status in LTC, which remains understudied in the CSCW community. Data work performed by them, we argue, goes beyond the "clerical and administrative work" [47] and functions also as a labor of maintenance and repair, that both guarantees data technologies' functionalities and acts as an articulation [68] between stakeholders. This new socio-technical insight about data work is crucial to data-driven collaborative LTC system design. Second, we deepen the understanding of data work by emphasizing its impact to data worker (frontline care workers in our study) and data-driven ecosystem. We reveal labor exploitation and increased unfairness are happening among these female, middle-aged, and not well-educated migrant workers, and is being normalized in name of data innovation, which should be taken seriously by not only healthcare researchers in CSCW, but also organizations and people with vested interests in data and data technologies. Third, our study informs design implications and recommendations for reducing care workers' burdensome data work, diminishing the negative impacts, and enabling more just, ethical and sustainable data-driven healthcare ecosystem.

2 RELATED WORK

Our study seeks to understand and detail "data-drivenness" in healthcare and broader scenarios from human participants' perspectives. Conceptually, "data-driven healthcare" in the paper refers to transforming numerous healthcare related aspects, actions and organizations into quantified data, and then using data and data technologies to drive healthcare practices and management. We define the scope of data as the combination of health-related data (for driving care planning and taking) and care-related data (for driving management). Health-related data are care receiver-centric, including physical, mental, behavioral, environmental, socioeconomic information pertinent to health and wellness. Care-related data include care plans, day-to-day care services, implementation, performance, and related administrative data. We define "data technology" as any technologies used to collect, store, process, and use these data types to drive care work and management. In this section, we first summarize prior research that outlines the promise and development of data-driven healthcare, then discuss studies that demonstrate the emerging data work in data-driven healthcare, and its practical impacts on care workers.

2.1 Data-Driven Healthcare: Promises and Challenges

The development of data-driven healthcare has been ongoing for decades. In the early stage, it was mainly directed at adopting and using health information technologies, such as patient portal [3] and electronic health records (EHR) [33], to support healthcare management. In recent years, the quick development of data sensing and computing technologies has brought disruptive changes to the term 'data-driven healthcare'. Specifically, various data-sensing technologies (e.g. sensors and mobile devices [46, 66]) have been designed to gather heterogeneous personal data known as "big data" in real time [5]. Synchronously, advanced data technologies (e.g. artificial intelligence and machine learning technologies [25]) are used to analyze and compute data, with the purposes of driving and supporting healthcare (e.g. developing personalized care plans [55, 76], assessing health [5], developing policy and making decisions [22], conducting research [24], etc.). This reshaped data-driven healthcare pattern provides tremendous promise for improving care performance, accountability, transparency, efficiency, and the cost-effectiveness of healthcare systems [60, 76].

The potential of data-driven healthcare is undoubtedly substantial. Under this pattern, the meaning and practices of ‘care work’ has been re-conceptualized beyond its typical use in the majority of medical literature [17, 34]. Data technologies also make the involvement and connection of multiple stakeholders, such as informal caregivers [35], volunteers [21], and patients themselves [65], much convenient. They form a complex care network and their communication and collaboration are mediated by data and data technologies [35].

Meanwhile, increasing researchers have reported on the problems and challenges of existing data-driven technologies in supporting practical care work. A large number of situated, personal data, for instance, have been recognized as playing significant roles in supporting practical care work, but usually lack in existing data systems [72]. Further, studies have found that many existing data-intensive infrastructures are established based on institutionalization models [61, 71], i.e. approaching caregiving practices as “work” consisting of a set of coded tasks, practices or problems to be solved. Yet real-world care-taking practices are situated, complex, and heterogeneous [10, 56, 72], which challenges the effectiveness of data-driven healthcare and even causes its cumbersome usability [8, 70], such as increasing caregivers’ burden [27, 72] and reducing labor productivity and efficiency [77]. Researchers in CSCW have articulated an important need to promote a dual vision of data-driven healthcare in which both its potentials and the challenges should be voiced. Our work contributes to this burgeoning stream of research by examining how data-driven LTC systems work in practice. While recognizing its benefits, we pay specific attention to its impacts on the care-taking practices of frontline care workers.

2.2 Data Work in Data-Driven Ecosystems

Data work, defined as “any human activity related to creating, collecting, managing, curating, analyzing, interpreting, and communicating data”[6], is usually thought a kind of articulation work [64, 68] that aids in develop and maintain data-driven infrastructure and making data-drive cooperative work work. Recently, a growing body of literature in CSCW and HCI has recognized the critical role of data work in various settings, such as data-driven healthcare [6] and scientific research [50], AI and machine learning models [62, 63], etc.

In the fields of healthcare, researchers have examined data work in different healthcare contexts, such as hospitals [47], community healthcare [69], emergency medical services [75], etc. Verdezoto et al., for instance, examine the invisible data work in community health in South India. They consider data work as a kind of articulation work, that helps bring together discontinuous elements in community health and manage the necessary actions to maintain community health infrastructures [69]. Bossen et al. [6] identify three major themes of data work in healthcare, including working with conflicting qualities of data, data work as collaborative endeavor, and relationships between patients and providers. Meanwhile, researchers have shown how people with traditional caregiving roles (e.g. doctor and nurse [6]) are involved into data work that far beyond past expectations, and how new occupations (e.g. medical scribes, medical secretaries, data analysts, data scientist, etc. [47]) have emerged to conduct data work in healthcare. Moller et al. [47] conduct a field study in hospitals in Denmark and U.S. and find despite the emergence of occupations like data analyst and data scientist, much of the work that makes data analysis, interpretation, and responsible use possible happens in administrative or clerical jobs.

Moving forward, researchers have examined the impacts of data work. Hoeyer et al., for instance, report that the datafication of healthcare brings a series of ‘meaningless data work’ to care workers in clinics mediated by digital technologies [28], and Normark et al. discuss the similar issue but used the term 'pseudo work' [32]. Particularly, many researchers have highlighted that data work is in and of itself a site where ethics and resistance play out [43]. They point out that while the role of data work is critical in, for instance, creating, maintaining, powering AI and machine learning systems [62, 63], and forming the infrastructure for technologies to function [69], much of this work is invisible, under-paid, under-valued and over-utilized [47]. Miceli and Posada [45] examine data annotators' data work for AI and machine learning models and argue that such data work entails the hierarchical valuation of labor and reproduces socio-economic power differentials that trap data workers (oftentimes in the "Global South") in precarious working conditions.

Despite the abovementioned studies, exiting efforts to data work are still in the initial stage, and a lot of data work that forms critical components of social and organizational infrastructure is missing from the discussion; meanwhile, much of the potential social significance and many of the consequences of data work necessitate further examination [6, 47]. Since these invisible data practices are growing exponentially in contemporary data-driven society, researchers in CSCW and HCI have articulated an urgent need to examine data work more broadly and more deeply in various health-care settings, which is crucial to elucidating the emerging requirements of data-driven system design needed to develop the workforce and skill mix necessary to carry out data work, organizational change, and the training of existing staff as well as to promote a more just and ethical data-driven future. Our work adds to this literature by specifically unpacking the data work of frontline care workers, which is largely underexplored in the literature.

2.3 Frontline Care Workers in Data-driven Healthcare

Issues related to frontline care worker have drawn significant attention in many related research fields. In CSCW and HCI, studies have started to explore the place of data technologies in supporting their caregiving practices, through, for instance, supporting communication [38] and collaboration [26] among care workers. Yet, most existing efforts have focused on supporting care workers' "care work". With the notable exceptions of [31, 69], little prior work has explored their increasing "data work" and how to support their "data work" in data-driven context. Further, a growing number of researchers have noted that most data-driven technologies primarily focus on the needs of care receivers and administrators that include designing sensing technologies (e.g. [36, 44]) to collect health information of care receivers and improve their healthcare outcomes [19, 59], and designing information storing and analyzing systems to aid care managers' decision making, control costs and improve efficiency [30, 41, 73], etc. The needs, conditions and data practices of care workers are largely underexplored [70]. This over-emphasizing patient/administrator-centric is challenging and may exacerbate already severe problems [27, 72]. Webster et al. [70], for instance, point out that while they do help workers better execute caregiving tasks or assist decision-making, the existing data-intensive healthcare systems are often hard to fix with care workers' highly context specific, fragmented and busy work routine.

Although these prior studies have identified the possible negative impacts of data-driven technologies to care workers, and suggested that the designing, developing, selecting and evaluating process of data-driven healthcare technologies should involve care workers [29, 48] and their care context [34, 35], most of their findings are general and lack of detailed reasons of these impacts. This causes the informed technology opportunities for addressing these issues lack of performability and specificity. In this paper, we deeply examined the practices, conditions, and challenges of frontline care workers through the lens of data work. We elaborated what and how the data work was performed by them and how this work impacted their current care-taking practices as well as the work of data-driven LTC. With data-drivenness becoming increasingly popular in healthcare, understanding these topics can shed light on how we can better design technologies to support care workers' practical work.

3 BACKGROUND AND RESEARCH SITE

The current study closely investigated data work performed by frontline care workers in a data-driven LTC facility in China, named izhaohu[2]. Because China has the largest and fastest aging population in the world, the topic of LTC for the elderly is extremely urgent [49, 57]. In response to this, Chinese government has taken 'technology mediated elderly care' as national strategies, and published a series of policies (e.g. [12, 51, 52]) to encourage the development of data-driven LTC. In line with national policies, more and more healthcare facilities in China have adopted and used data-intensive technologies to transform healthcare work and management. Izhaohu ('爱照护' in Chinese, in which '爱' means 'Love', '照护' means 'care', and '爱照护' means 'caring with love') is a popular one among them.

The development philosophy of izhaohu is building plenty of what are known as IACC (Intelligentized All-Inclusive Care Community) and providing intelligent LTC for the elderly in China. The word "intelligent" is izhaohu's core and refers to its use of various data technologies and products to gather as much of seniors' health and care data as possible, then use those data to connect every element of their care in order to support various caregiving practices among care participants and drive and manage the overall caregiving work. As opposed to the well-known continuing care retirement community model [74], IACC is deeply rooted in the traditional Chinese family value of aging at home [54] and is usually small and embedded in local communities. Because of its consideration of China's cultural background and taking full advantage of various digital technologies, izhaohu has quickly become one of the most well-known LTC institutions in China.

The context of izhaohu allowed us to closely examine several unique dimensions of data work in data-driven health care. First, to support the concept of IACC, a hybrid digital infrastructure has been implemented in izhaohu, ranging from health data collecting and tracking, storing, computing devices, to communications and management tools (some of them were shown in Fig. 1). During daily caregiving, an overwhelmingly large amount of health and behavior data were sensed and captured by various data-sensing devices (e.g. Figures 1-a to 1-d) and integrated in intelligent workstations (Fig. 1-e). These and other user data consisted of senior's Electronic Caregiving Records (ECRs, figures 1-g and 1-h). In ECR system, many AI algorithms were coded to analyze and compute the data, which were then used to drive and support care participants' caregiving work, including tasks related to real-time emergency warning for care staff (Fig. 1-f) as well as remote monitoring and management for care managers (Fig. 1-g) and families (Fig. 1-h). Thus, the facility was a valuable site for investigating how these data and data technologies work in a real-world LTC context. Second, while the IACC model offers seniors a familiar living environment, it also provides more convenience for various informal caregivers (e.g. families, neighbors, and old friends) to collaboratively care for seniors. We therefore gained the opportunity to learn how data and data technologies support communication and collaboration among various stakeholders, which is crucial for understanding how data and data technologies work.

---

[2] http://www.izhaohu.com

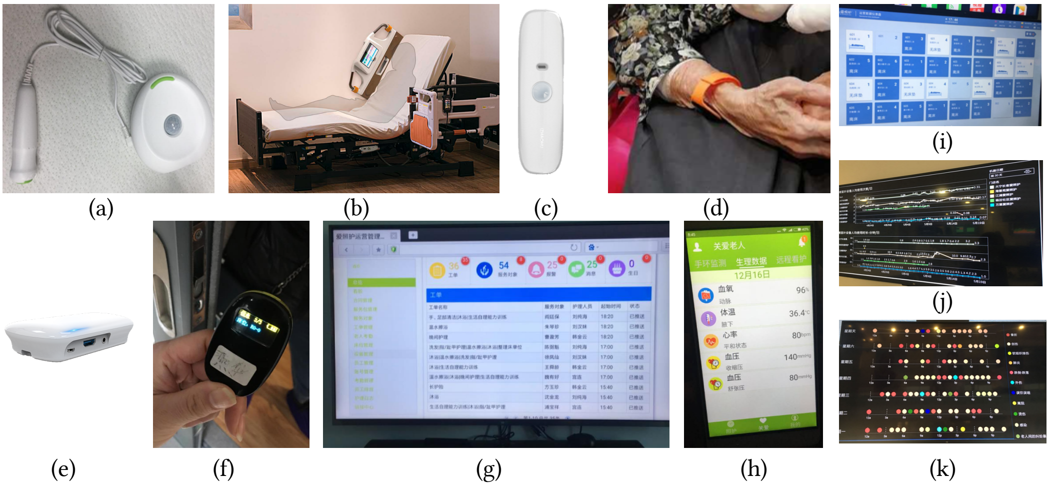


Fig. 1. Technologies used in izhaohu: (a) an infrared sensor for monitoring seniors' actions and alerting care workers when they fall down; (b) a smart bed equipped with several sleep-monitoring sensors; (c) a sensor for monitoring seniors' defecation condition; (d) a smart band for monitoring seniors' location and vital signs; (e) an intelligent workstation for storing, processing, and submitting data; (f) "little mouse" receiving the alert sent by sensors or systems; (g) electronic caregiving record web terminal used by managers; (h) electronic caregiving record mobile terminal used by families; and (i-k) three data dashboards.

4 METHODS

We adopted qualitative research methods, including participant offline observations and interview studies, to explore our research questions.

4.1 Data Collection: Participatory Observation

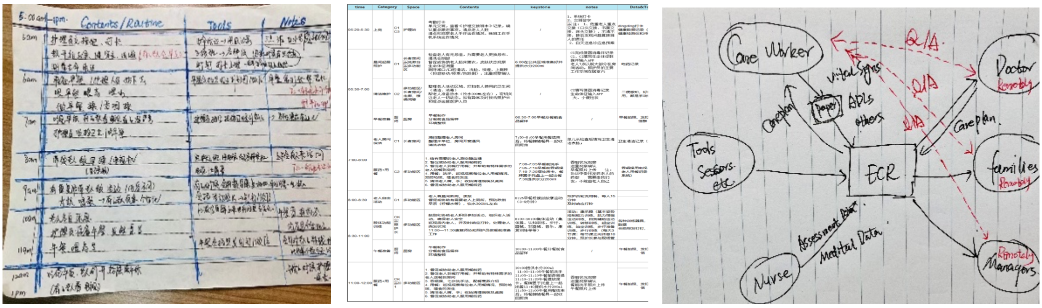
We began with a three-week participatory observation with the purposes of obtaining a general understanding of data-driven care work in izhaohu (RQ1) and shaping the subsequent interviews. We contacted izhaohu through its call for cooperation[3] and explained our research purpose, requirements (including long-term participatory observations and interviews with people), and commitments to data privacy and safety. The manager in izhaohu was very supportive of our research and appreciated our interests in them. They gave us permission to conduct our study in one of their care facilities in Shanghai China.

4.1.1 Data Collection. On the first day, izhaohu's manager accompanied us to the facility and introduced us the overall situation. Next, the first author and two assistants, three native Mandarin speakers, started the three-week observation. To gain its complete process, we three observed in three shifts each day (5 a.m.–1 p.m., 1 p.m.–9 p.m., and 9 p.m.–5 a.m.). The space design of izhaohu was family style. Our observations were mainly conducted in the reception room, living room, dining room, and practice room.

The whole process was divided into two stages. The initial stage, lasting two weeks, was for building the broad understanding of the situation at izhaohu. We observed and recorded the daily caregiving process, including people, technologies, practices and activities. We were careful never to disrupt these processes, but we also interacted with people, followed various care cases, and tracked critical incidents occurring randomly.

---

[3] http://izhaohu.com/#/cooperation



(a) (b) (c)

Fig. 2. Examples of observational data: (a) a handwritten field note in stage I; (b) shared document; (c) a field note in stage II.

During this stage, we mainly collected data in the form of a daily schedule, including time, practices, details, tools, stakeholders, and notes. We also gathered relevant organizational and management documents. We used handwritten field notes (e.g. Fig. 2-a), photographs, audios, and videos to collect data. We built a shared document on Google Drive to collaboratively edit the daily schedule and gather other data (Fig. 2-b). We gathered and reviewed the recorded data, and filled in/updated the shared document every two days. After this stage, we identified izhaohu's general caregiving routine and workflows, the main data technologies (parts of them shown in Fig. 1), and the main stakeholders and their general practices and activities: frontline care workers who conducted the major daily caregiving work; medical caregivers (e.g., doctors, nurses) who took charge of assessing seniors' conditions, making a care plan, and providing basic medical care services; and informal caregivers (e.g., families, friends) who frequently visited seniors and provided necessary personal, and emotional care.

In the second stage, lasting one week (24 hours per day), we mainly focused on izhaohu's data-loop and data-related practices and activities (i.e., which data were collected and recorded, for what and how, the tools and technologies used, how data were used to drive daily health care, and how care participants used data to reach their goals). We also watched for new phenomena. Meanwhile, the first author began to establish rapport with people and was subsequently invited to a series of formal and informal offline activities and events, including family caregiver meetings, care staff's daily scrums, case discussion meetings, and design discussions related to data products. Observing these activities and engaging in interactions with involved people helped us gain deeper insights into different means of using and analyzing data. We still collected data in the form of handwritten field notes, photographs, audios, and videos. Additionally, we used sketches to help with mapping and visualizing the interactivities among data, data technologies, and stakeholders (e.g., Fig. 2-c) and to facilitate reflections on the existing data practices and challenges of frontline care workers.

Through three-week participatory observation, we got the preliminary answer of *how data and data technologies work in the LTC* (RQ1) and *what and how is the data work performed by frontline care workers in the contexts of data-driven long-term care* (RQ2).

4.2 Data Collection: Semi-Structure Interviews

To address the question of *How does this data work introduced by 'data-drivenness' impact frontline care workers' practical care-taking work* (RQ3) and to help address certain questions that we could not during participant observations for RQ1 and RQ2, we conducted in-depth interviews with frontline care workers and other involved stakeholders in izhaohu.

4.2.1 Recruitment and Criteria. We recruited care workers using two approaches: 17 were recruited during participant-observation, and two were recruited during our follow-up tracking of the happened events. Using purposive sampling, we actively interacted with care workers who performed and were engaged in a lot of data work throughout participant observation. We disclosed our intention and recruited them for our interviews. The remaining two care workers were the stakeholders of specific events that occurred during participant-observation. For one event, the dealing process often engaged a lot of stakeholders. When we tracked specific events, we actively reached stakeholders who were engaged in data-related interactivities, then recruited the remaining two participants. Of the 19 care workers, 17 were female and 2 were male; their ages ranged from 38 to 57 (M = 46.7 years old; SD = 5.1 years). Their education backgrounds were generally not very high (one primary school, three junior high school, eight senior high school, five technical school, two junior college).

Additionally, we also interviewed 13 involved stakeholders (five medical caregivers and eight families) as supplementary data. We focused only on the data practices they were engaged when using data and data technologies to reach their own caretaking goals and their interactions with care workers, if any. Two medical caregivers and six families were recruited during the participatory observation, and the others were recruited when we followed up on specific events. These interviews functioned as supplementary data and provided multiple insights into data interactions between other stakeholders and care workers.

4.2.2 Data Collection. In our interviews with 19 frontline care workers, we asked about details of 1) how data and data technologies specifically impacted their daily care work (benefits and inconveniences); 2) their reactions toward inconvenient situations, if any; and 3) their experiences and perceptions of data work and data-driven health care. These questions were designed to be general so that they could be inclusive and encourage interviewees to share various levels of practices, experiences, and perceptions. When interviewees mentioned some interesting points or prior experiences, we followed up asking for more details and concrete examples. Meanwhile, we also had a range of tailored questions targeting each specific interviewee based on participatory observation. In the interviews with 13 stakeholders, we mainly asked about their interactions with care workers, data practices they were engaged in when using data to reach their own caretaking goals, and their experiences with data-driven health care.

The first author conducted all the interviews face to face and in Mandarin in izhaohu. Each lasted for 40-120 minutes. During the interviews, with participants' permission, we also obtained other forms of data to triangulate their interview responses, including 1) screenshots of their mobile app pages and 2) related photos or videos they shot, if any. When some participants did not understand the terms "data," "data technology," or "data work" very well, we used simple examples collected during participatory observation to illustrate the vocabulary. All interviews were audio recorded and transcribed into Chinese verbatim later.

4.3 Data Analysis

The analysis was based on the inductive process of grounded theory [23]. The first and second authors started coding and analysis while data were collected. During the open coding phase, we independently reviewed the collected data back and forth, and generated codes and materials (e.g. sketches) to identify emerging themes. We regularly conferred with each other to ensure reliability, and reviewed and verified these codes and materials as new data were obtained. After this stage, we generated an initial set of codes, capturing frontline care workers' data practices, experiences, concerns, encounters and expectations related to data-driven healthcare. Several preliminary themes were identified, such as the reconfigured care work (care workers' role and responsibility, workplace, workflow, work performance), the roles and functions of data technologies, the tensions between what care participants expected and what existing technologies provided (data contents, transparency and accountability), encountered data-related benefits and inconvenience, frontline care workers' formal and informal data practices (gathering, recording, sharing, interpreting, contextualizing, checking, communicating), potential ethical issues (workplace discrimination, fairness), etc.

Based on these preliminary themes, we re-focused our analysis to the broader level of themes and tried to consolidate our codes into one overarching theme. We iteratively sorted different codes into potential overarching themes and collated all the related data within the identified themes. Through a few iterations, we identified our final thematic map: three primary themes concerning three distinct types of data work frontline care workers conducted to make data-driven healthcare work. In the following sections, we present the full details of these themes. We use representative quotes, which were translated into English for this paper to illustrate our points. For protecting participants' identities, we used C1-C19 to denote 19 frontline care workers, M1-M5 to denote 5 medical caregivers (2 doctors and 3 nurses), and F1-F8 to denote 8 families. The other names used in this paper are pseudonyms.

4.4 Ethical Considerations

All three researchers received the formal ethics training, and were granted official certificates from ethics committee in their institute. Before starting, we informed all care workers and seniors in the facility our intention and background information, and got their permission. When we collected their individual data (e.g., screenshots of their mobile phones), we asked for their permission again. We promised all collected data would be used only in this study. All collected data were used in anonymized ways (i.e., there was no link between the collected data and an individual user). Our institute does not fall under the jurisdiction of an internal review board or similar ethics committee. We did, however, get feedback from peers inside and outside our faculty to validate the ethical perspective of our research.

5 FINDINGS

Our findings focus on frontline care workers' data practices, roles, responsibilities, experiences encountered challenges as data workers. Before providing specific details, we will first outline data-driven LTC in izhaohu – the integrated data and data-technologies and the reconfigured data-driven healthcare. Next, we will unpack the practices of data work performed by frontline care workers, and how these practices function as a labor of maintenance and repair, that both guarantees data technologies' functionalities and acts as an interface between stakeholders (e.g. medical caregivers and families) and data systems.

5.1 Data-driven Long-Term Care

Given the development concept of 'data-drivenness', the main stakeholders in izhaohu – care receivers (mostly seniors) and care providers (mostly frontline care workers) – were described as "pixelated persons" with specific digital user portraits, which were the foundation of data-driven

LTC. We inducted data-driven LTC in izhaohu into two categories: elderly-centric data-driven care and care-worker-centric data-driven management.

Table 1. Data and data technologies involved in izhaohu

Data		**Data Technologies**	
Category	Main Contents	Category	Technologies
Health and wellbeing data (comprehensive)	· Individual profile · Preference · Physical condition · Cognitive condition · Mental condition · Social relationship · Disease history	Data sensing	· Smart cameras · Infrared sensor (Fig. 1-a) · Smart bed (Fig. 1-b) · Urinary and fecal sensing device (Fig. 1-c) · Behavior sensing device · Smart band (Fig. 1-d)
Health and wellbeing data (daily)	· Daily vital signs · Activities of daily living · Medical records · Abnormal conditions · Visiting/outing records	Data storing, computing	· Data workstation (Fig. 1-e) · AI
Care data	· Care plan · Care plan execution · Care events	Data applying	· Alarming application (Fig. 1-e) · ECR (Fig. 1-g, 1-h) · Data visual dashboard (e.g. Fig. 1-i, j, k) · Office Automation (OA) system
Administrative Data	· Attendance data · Performance data · others		

5.1.1 Date-Driven Care-Taking. In data-driven izhaohu, seniors were described by a digital ECR, which were established on the first day the senior lived in the facility and constantly updated on each of the following days. Described by M5 (a nurse), the data in ECR were extremely rich (more than 300 items, listed on 22 sheets) and multidimensional (see the details in Table 1). We induced these data into three primary categories. The first was comprehensive health/well-being data, including seniors' individual profiles, preferences, social relationships, disease histories, and physical, cognitive, and mental conditions, social relationship and functions, disease history. These data were collaboratively obtained after seniors lived in izhaohu and updated each month by nurses and frontline care workers. The second category was daily health, well-being, and behavioral data, including seniors' daily vital signs (e.g., body temperature, blood pressure, heartbeat, and blood oxygen), ADLs (e.g., eating, bathing, sleeping, getting dressed, and toileting), medical records, abnormal conditions, and other data (e.g., visiting records). These data were collected and recorded by frontline care workers every day. The third category was care planning and executing data. Based on the comprehensive health and well-being data, a specific care plan for each senior was collaboratively generated by nurses, doctors, and embedded smart algorithms. Frontline care workers provided care services based on this plan. The execution process and random events were also recorded.

The involved data technologies primarily (see details in Table 1) fell into three categories: 1) data sensing technologies (e.g., cameras, sensors, smart bands), which were mainly used to collect various health and behavioral data; 2) data storing and processing technologies, which were embedded in the ECR system for analyzing and computing data such as assessing seniors' health-care condition, planning care services, and evaluating care performance; and 3) data-applying devices (e.g., alarm tools for care workers, portals for different stakeholders, a data-visualization dashboard, and management platforms for managers).

There were three main stakeholders. The first were frontline care workers who stayed in the facilities and conducted the major day-to-day caregiving work. Most were female (88.9%), middle-aged (the average age was 49.1), and undereducated migrant workers (100% in izhaohu). They were also the primary holders and users of many data technologies. The second category were medical caregivers, including doctors, nurses, and therapists. The took charge of assessing seniors' clinical conditions, making care plans, and providing basic medical care services. They stayed in the facility on a fixed day every week. Other times, they remotely learned about and monitored seniors' conditions via ECR medical clients and provided onsite medical care services when needed. The third category was informal caregivers, mainly seniors' families or other guardians. They remotely learned about seniors' conditions via ECR family clients, visited seniors, and provided necessary personal, social, and emotional care when needed. Other roles included care managers and technical engineers.

Fig. 3 shows a brief view of the data-driven caretaking process at izhaohu. We reduced it to three subprocesses. When one senior applied to live in izhaohu, a nurse conducted a preliminary judgment of whether the senior could live there based on some basic standards (e.g., "*are there any infectious diseases, care risk, serious illness, etc.?*" [M5]). If the senior was permitted, they would begin an onboarding process, and a personal ECR account would be created. Then nurses and care workers (usually the group leaders of frontline care workers) would conduct a comprehensive health assessment, and the comprehensive health/well-being data would be collected and recorded. Nurses and care workers would collaboratively record the information into the senior's personal ECR account. After that, the embedded algorithms in the ECR would be triggered to analyze the senior's health conditions and generate specific care program. The program would be sent to doctors and families through their ECR client for optimizing and confirming. After the final care program was generated, the workflow would enter into the care-delivering process in which frontline care workers would check the senior's care program through their ECR app and deliver care services to the senior accordingly. Each day, daily health, well-being, and behavioral data would be collected and recorded into the senior's personal ECR by care workers and various sensors collaboratively. Meanwhile, medical care workers (e.g., doctors and nurses) and families could remotely learn about seniors' conditions, intervene in a timely manner, or adjust care programs through their ECR client (Fig. 1-g). At the end of each month, the monthly care-evaluating process would be triggered, and nurses and care workers would conduct the assessment work again and update seniors' health assessment result. The recorded data were used to assess seniors' health conditions as well as care workers' work performance.

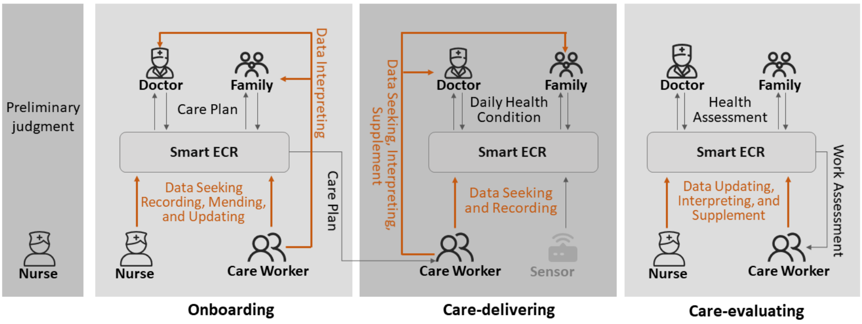


Fig. 3. Overview of data-driven care-taking process in izhaohu

5.1.2 Date-Driven Management. "Data-driven management" in izhaohu referred to using data and data technologies to monitor caretaking process and evaluate care workers' work performance. The data included caretaking records (e.g., care plan execution, response time and rate to system alerts, number of human-caused accidents, and complaint rate), number of performed tasks, and attendance records. The data technologies included data-collecting tools (e.g., cameras) and managing tools (e.g., data visual dashboard and OA systems).

The whole process described above seemed reasonable and actionable, and data-intensive ECR systems worked as the pivot to drive everything along. Yet our analysis suggested frontline care workers were actually the crux of this data-driven configuration, and they conducted a series of data work [6] to underpin and implement the facility's data-driven vision, including constructing, maintaining and updating data infrastructures, articulating the distributed and fragmented data and making data-driven multiple collaboration work, and repairing the breakdown of data infrastructures. The red arrows in Fig. 3 list some of this data work. In the next three sections, we turn our attention to frontline care workers' data work and elaborate these three kinds of data work they perform and how it impacts frontline care workers' practical care work.

5.2 Data Work to Construct and Maintain Data Infrastructures

Fig. 4. Different forms of data and data work in izhaohu. a) A slogan posted in izhaohu that reads, "If we have data, please follow the decision of data. If not, please follow the decision of CEO"; b) a care worker transferring data from her own notebook onto a formal form; c) the interface through which care staff recorded daily caregiving data online (the left column shows the data category, including vital signs, mental condition, nutrition, skin, ADL, sleep, pain, medicine, and blood glucose, and the left column shows the details of nutrition); d) a screenshot of a WeChat group wherein a daughter was asking for her mother's blood pressure; e) the photo a care staff member took to record a happy moment experienced by the senior she cared for; f) a care staff member's notebook in which seniors' detailed conditions were recorded; g) display board and printed photos used to show care staff's care work; and (h) a paper form recording a senior's daily vital signs and ADL.

The implementation of data-driven healthcare requires vast amounts of data. While many data-sensing technologies (shown in Fig. 1) were used to sense, collect and store data, the data they could sense were limited, only including part of vital signs and ADLs (e.g., sleeping, walking, and water-drinking data). Most of data required in ECR and work management systems were hard to be collected automatically. For data-drivenness, frontline care workers, who knew the most about the seniors and care work, therefore became the actual data producers and maintainers.

5.2.1 Constructing and Maintaining ECR Data. The data in ECR system were the core of data-driven LTC. In izhaohu, most of these data were manually produced, collected, and recorded by frontline care workers. The comprehensive health and well-being data, for instance, were collected and recorded by care workers during the onboarding process. As described in section 5.1, when a senior applied to live in izhaohu, care workers would interview care receivers and their families with nurses and capture and record all the assessment data on 22 sheets (nurses were in charge of medical data; care workers were in charge of care data). Often, they recorded all data on a paper form during interviews and transcribed them digitally later, even though an iPad application (Fig. 4-c) was specifically designed for them to record data directly into systems. According to C4, "*The iPad is not very convenient when interviewing with seniors.*"

The daily health, wellbeing and behavioral data were recorded by frontline care workers in day-to-day care-delivering process. Every day, care workers needed to record several bits of information in both paper forms (e.g. Fig. 4-b) and online forms (e.g. Fig. 4-c). Meanwhile, due to the inherent personalized, situated and dynamic nature [10] of the caregiving process, many healthcare-related data were not static, and needed to be updated frequently. This updating work were also performed by frontline care workers.

This data work was complex, labor intensive, and time consuming, because izhaohu's data-driven vision needed as much and as detailed data as possible. The situated and complex nature of LTC for the elderly further made some of the required data very hard to accurately gather, format, and record in real time. Most of our participants perceived data work as a "*heavy burden*" that exacerbated their already busy and stressful care work. They complained that their daily work was "*strongly surrounded by various data and forms, and struggled in heavy and complex data work every day*" [C7]. Further, our participants expressed they had almost no time to do this data work during working hours, therefore they often performed that in their rest or off-duty time. As C13 said:

> *"We are required to record so much data. But how can this happen? Imagine, for example, eating. The required data includes how much staple food, how much soup, how much meat, how many vegetables, eating time, how long they spend eating, whether they eat independently and so on. There are dozens of old adults here. If we record this data immediately when it happens, we can't do any other things." [C13]*

Particularly, this additional data work was considered to be care workers' default work. Instead of providing additional compensation, care workers' work performance would be impacted if they didn't record these data as required. These features significantly impacted care workers' positivity and attitudes toward data work, which in turn impacted the quality of recorded data. In the recording process, we found some care workers took down easily forgotten data (e.g., vital signs) and special events in their own notebooks, then transcribed them on form paper (Fig. 4-b) and in the online data system (Fig. 4-c) during their off-hours. But not every care worker did that. More often, they produced data through recalling the care work during on-hours or by improvising, which greatly affected the data quality. For instance, in one paper document of one senior's vital signs and ADL (Fig. 4-h), the recorded contents were almost same: the sleeping quality was "general," the senior was "eating everything provided in three meals," and the amount of drinking water was the same. This led to what Sambasivan et al. defined as "data cascades" [62] and caused negative, downstream effects to data-driven healthcare.

5.2.2 Constructing and Maintaining Work Data. In data-driven context, frontline care workers' job performance was evaluated and calculated based on data, including attendance time, response timeliness to alerts, accident rates, etc. While some of these data could be collected through sensing technologies (e.g. collecting attendance time through workers' network signals), most of them depended on care workers' manual recording. For instance, care workers needed to manually punch the clock at required times or in required places to ensure they finished the specific task as required. Although these procedures were helpful in driving managerial processes, our participants thought that the procedures greatly increased their workload and made them feel like "*being controlled by data* [C7]." For example, when an alarm was activated, care workers were required to check it in five minutes, or else *"the system will label it as 'red', i.e. we don't deal with it in time"* [C3].

Meanwhile, due to the inherently complex and uncertain nature of LTC [10], much of work care workers undertook was invisible [47] and hard to accurately quantify and record in the standardized evaluating system, which caused the often misrepresentation of their workload. Consequently, frontline care workers had to use notebooks or photos to record what they had done, and provide to managers as a kind of workload record.

In summary, data-driven LTC is highly dependent on various types of data. Yet, existing data technologies couldn't sense and record most of them very well. Consequently, frontline care workers became the practical data recorders and maintainers. They collected the required health-, care- and work-related data and inputted into data systems. These data were the key component of data infrastructures and were used to drive the subsequent care-taking and management practices. From the perspective of frontline care workers, however, this work was very onerous and worsened their already busy and stressful care work, driving them to exhaustion.

5.3 Data Work as an Articulation Between Stakeholders

The LTC process for the elderly depended on the collaborations among several stakeholders, especially frontline care workers, medical care workers, and families. In izhaohu, medical care workers and families mainly learned seniors' situations remotely through their own ECR clients, and provided the necessary support and services when needed. Frontline care workers, as the people who conducted the most day-to-day care work, often worked as an articulation to gather and mesh multiple demands and opinions and make the data-driven multiple collaboration work.

Taking the generating process of individual care program as an example, as mentioned in section 5.1, the care programs were synergistically determined by frontline care workers, medical care providers (e.g., nurses and doctors), and families. This process seemed ECR driven – care workers and nurses evaluated seniors' health and well-being conditions, generated the preliminary plan, and submitted it to ECR; doctors checked and confirmed the information via their ECR clients and submitted again; then families checked, confirmed, and generated the final version of care program. Yet what we observed in practice was a frontline care workers driven process. Specifically, each stakeholder had its own demand: medical care providers focused more on clinical features and services; frontline care workers focused more on daily habits and preferences; and families considered things like economic factors. Frontline care workers, then, became articulations to communicate, coordinate, and push the generation of care programs. As C3 explained,

> "*The care program proposed by doctors is usually very professional. But what families consider is usually more comprehensive. They will consider their family financial situation, family relationship, etc. They usually require us to modify some contents because 'it was too*

> *expensive' or there was 'no need'. Under the circumstances, we need to supply these factors they consider into the system, discuss them with doctors, and adjust care program through consultation.*" [C3]

Moving forward, during the daily caretaking process, the elderly's demands frequently changed due to various reasons such as the change of health conditions or family relationship. The changes of individual demands brought corresponding changes of care programs. Frontline care workers then needed to communicate seniors' (or families') requirements to care managers, negotiate the development of a new program, and update the ECR system. This process was time consuming, and kept happening throughout the whole caregiving process. C19 provided the example of Grandpa L to explain.

> "*After the first assessment, we suggested a detailed care plan. His daughter accepted it, but a few days later asked to cancel many rehabilitation training sessions because Grandpa L felt very tired. Later, his daughter asked us to add these sessions back because she found Grandpa L always sitting there, not getting any exercise. She doesn't want that. We then update the data once again.*" [C19]

Additionally, frontline care workers also needed to report seniors' situations to care managers and families, answer families' questions, respond their requirements, etc. While our findings in the section demonstrated the significant role of frontline care workers in gathering and meshing multiple demands and opinions, and making the data-driven collaboration work, these also demonstrated the fails of data-intensive system in articulating this collaborative LTC work. As the main care providers, frontline care workers then had to repair these fails. In the next section, we will elaborate what type of data work was frontline care workers performed to repair the breakdowns and how they did it.

5.4 Data Work for Repairing the Breakdowns of Data Infrastructures

Previous studies have shown that LTC for the elderly is inherently complex, situated, uncertain, and fragmented, which greatly challenged the utility and efficiency of data technologies in LTC settings. In our study, we have identified a series of breakdowns of data infrastructures in driving caretaking and management, due to, for instance, collecting incomplete or incorrect data, missing useful and valuable data, etc. For addressing these, a series of corresponding data practices were performed by frontline care workers to repair these breakdowns and make data-driven LTC work.

5.4.1 Repairing Data Errors and Incompleteness: Restoring and Supplementing Data. As illustrated in section 5.1, izhaohu had a series of sensory devices helping practitioners sense and collect seniors' health and behavioral data. Yet accurately sensing health-related data had always been a big challenge for sensory technologies. In the context of elderly caregiving, seniors' huge differences and diversity as well as the situated and uncertain nature of care work, further aggravated this challenge. Incomplete or incorrect data were unavoidable. Under the circumstances, frontline care workers, becoming data restorers, performed data-correcting and supplementing practices to repair incorrect or incomplete data.

Smart bed (Fig. 1-b), for instance, was a typical sensory technology in izhaohu. Based on its specifications, several sensors, including a pressure-sensing device, a sleep-monitoring device, a bedsore-perceiving device, were embedded into the mattress to gather data around seniors' sleep. The sensed data were analyzed and computed in real time. If an exception occurred, an alert was sent out. For instance, the pressure-sensing device was used to sense whether seniors stayed in bed throughout the night sleeping time. If not, an alert was sent out to the receiving terminal, named *Little Mouse* (Fig. 1-f), which reminded care workers to check senior's situation. Synchronously, the ECR system and data visual dashboard (Fig. 1-i) recorded and showed the alert, and stakeholders (e.g., care managers and families) also could check this data through their own ECR clients. Yet the seemingly reasonable data-drive logic was usually broken by error data. Often, the alert was wrongly sent out when, for instance, a lighter-weight senior moved to the edge of the bed or a senior accidentally touched the alarm button. Frontline care workers then needed to manually correct the data and interpret them to stakeholders when needed. F5 was the daughter of Grandma M, who suffered from serious dementia and had lived in izhaohu for more than two years. She had experienced such an 'data accident':

> "*I remember one night the alarm number send from my mother's bed was very high. When I checked the app in the next morning, I was scared and didn't know what happened to my mom. I immediately called Jia (care worker who was on duty that night) to ask. You guess what, she told me my mom considered the call button as a new toy that night and was pressing it for fun.*" [F5]

In the second day, *Jia* repaired these 'wrong' data and make the ECR normal. Besides the incorrect data, the situated and uncertain nature of LTC often led to the incomplete data. But under the institutionalized data-driven context, this incompleteness was not allowed. Frontline care workers thus needed to supplement the missing data. For instance, to strengthen the management of caretaking at night, izhaohu required night-shift care workers to patrol every two hours (i.e., 22 p.m., 24 p.m., 2 a.m., and 4 a.m.). The sensors in the seniors' smart bed collected patrolling data by sensing the distance between bed and care workers' *little mouse*. Yet this seemingly plausible data-driven management process was also usually failed in practice because some seniors often got up before 4 a.m., which caused the incompleteness of data. Care workers then need to label and explain this incompleteness.

5.4.2 Repairing the Breakdown of Data Infrastructures in Evaluating Seniors: Collecting and Recording Meaningful Data. One significant role of recorded health data was evaluating seniors' conditions, and supporting the following care-taking and decision-making. However, evaluating a frail person was different from evaluating a patient with a specific disease, because the disease often had clear evaluation indicators, but a frail person didn't have. This greatly challenged the utility and efficiency of data-driven health management LTC settings. In izhaohu, although the ECR system recorded a large amount of senior-centric data, our findings suggested these data were still insufficient in clearly describing a person and driving care-taking and decision-making. Most stakeholders (e.g. families and medical care providers) expressed they couldn't learn the conditions of seniors very well through data in ECR systems.

What we found in addition was that frontline care workers often used their own methods to record what they considered meaningful and valuable data. The recording forms were diverse, such as taking photographs, shooting videos, writing diaries, etc. Fig. 4-e, for instance, taken by C5, recorded the happy mood of a senior. Fig. 4-f was C12's diary, recording Grandma Ye's conditions:

> *"Grandma Ye has dysphrenia. At the beginning, she didn't eat. Her daughter told me she hadn't eaten normally for a long time (before living in the facility). We then thought about and discussed what to do. We remembered her daughter told us she was Christian. We thought this might be a solution. At mealtime, we played Christian songs and told her the Lord gave the food to us. It works!...We are so proud of Grandma's change. Currently, she can eat by herself. We've improved her life quality greatly!"* [C12's diary]

Specifically, compared to negative attitudes toward stressful formal data recording work described in section 5.2, this informal data recording practice was voluntary, and our participants expressed high initiative and enthusiasm to do that. Throughout our interactions and conversations with frontline care workers, they often took out their mobile phones to show us seniors' tiny moments they recorded, such as seniors' eating, playing, exercising, etc. There were also some continuous records which showed the remarkable improvements of seniors, such as from "*unable to eat independently*" to "*delighted to eat by himself,*" "*unable to walk independently*" to *"able to go to toilet independently.*" Uncle Wu's story, recorded by C2's in photos, was a typical one. In our interview with C2, she excitedly showed us her photos which recorded the impressive improvement of Uncle Wu:

> "*You see, this was Uncle Wu when he just came here. This is what he looks like now. He looks very great, right? He's eating by himself! At the beginning, Uncle Wu didn't talk to anybody and wouldn't eat. We used our methods to care for him...Two weeks later, his son came to visit him, and he called out his son's name. His son was immediately moved to tears because Uncle Wu had not known his son for more than two years.*" [C2]

This recorded informal data were owned to care workers themselves, but often used in the formal occasions, such as their morning conference and weekly conference to present seniors' condition. Care workers also printed and posted some of photographs in a public space (e.g. Fig. 4-g) to show some touching and happy moments. As opposed to the standardized and institutionalized data recorded in the ECR, our participants thought these more nuanced details represented the truest and most substantial experiences of seniors and should have been added to seniors' individual electronic profiles. Yet they were universally absent in the ECR system, which participants a "*a pity*" and "*a little unreasonable.*" As C7 said,

> "*If only focusing on medical data, you can never get a positive result. Seniors here are too old. From a medical perspective, their physical conditions must be continually declining, no matter how we care them. But many seniors do show some improvements. Doctors here gave us very big recognition and considered these improvements as medical miracles. Yet the current data system is only based on those physical data and ignores these details.*" [C7]

5.4.3 Repairing the Breakdown of Data Infrastructures in Supporting Daily Care-taking Work: Informal Data Seeking, Exchanging, and Discussing. As for the roles and functions of data infrastructures in supporting day-to-day care work, the main work contents of frontline care workers, our participants (frontline care workers) talked about it as "*useful, but limited*". They basically agreed that the enormous and heterogeneous digital profiles in ECR could helped them learn seniors' basic conditions at the beginning, but didn't function very well in supporting daily care work. They felt that they knew the seniors better than the data. As C6, a very experienced care worker, explained: "*We stayed with seniors all day and already knew their daily conditions very well. There was no need to learn seniors' conditions through those data.*" Even though the ECR systems contained some expert algorithms that could generate some professional analysis, most of our participants thought it played what C1 called "*a limited role*" because "*professional knowledge was usually taken out of practical context and was hard to use.* [C1]"

Our participants emphasized that caring for people in izhaohu required not only standardized health and well-being data recorded in ECR but also various social, emotional, cultural, and family data. The latter was even more important for improving the quality of care, but they could not obtain such data from ECR. C13 told us that when she got a so-call professional care plan from ECR, she usually did not know how to care due to the lack of key personalized conditions of seniors, such as "*what foods and clothes they like and what they don't like, their hobbies, their experiences, the specific attention points, etc.*" She considered these were important for them to care for seniors, but always absent in ECR. She raised the case of Grandma C, who suffered from serious dementia, usually did not eat anything, and was often in a bad mood, to explain why those data were important:

> "*Imagine, if Grandma C doesn't eat anything, what can you do? I don't think that that heterogeneous and very professional data could give me the answer. No doctor can tell you how to ask a senior to eat, even if they are on the spot. But if we know she has a lovely grandson who is in elementary school and she loves this grandson very much, we then say something like 'After eating, your grandson will be very happy and lover you more' to coax her. It usually works very well.*" [C13]

Similar to Grandma C's case, our participants showed us many cases in which they used various individual and lifestyle data – such as seniors' preferences, family relationships, psychosocial and socio-emotional characteristics, and day-to-day behavior – to cope with daily care problems. "*Grandpa Xi once had a boss named LaoWang,*" said C5. "*If he doesn't eat, we usually say, 'LaoWang asked you to eat. If you don't, he will take some money out of you monthly earnings.'*" And C14 said, "*Grandma Li was Christian. We usually play Christian music to her to make her happy.*" Our participants further expressed these daily care problems, such as seniors refused to eat or be in bad moods, happened very frequently and were actually the most important things for caring the frail seniors. For coping with these problems, they thus conducted various informal data practices – data seeking, exchanging and discussing – to obtain the information they needed. For example, during the first day a senior came to izhaohu, care workers often conversated with the family to obtain those data. These data exchanging practices were informal, voluntary, regular occurrence and continued throughout seniors' residence in the facility.

Besides informal data seeking and exchanging practices with families, frontline care workers themselves also often exchanged and discussed data, and analyzed seniors' health, well-being, and behavioral information before developing their own care methods. Exchanging and discussing data were the most common informal data practices frontline care workers engaged in. They went about these practices in a variety of forms, including corridor talk, daily scrums, and weekly meeting. During these processes, the information about seniors' specific status, behaviors, characteristics, and problems were discussed in depth. Both formal data in the ECR system and informal data recorded by care workers themselves were used as auxiliary evidence to understand seniors' conditions. Most participants said these informal data practices were more helpful when it came to better caring for seniors. Similar to informal data recording practices in section 5.4.2, these informal data discussion practices were also performed by care workers proactively.

5.4.4 Repairing the Breakdown of Data Infrastructures in Meeting Stakeholders' Needs: Data Supplementing, Interpreting and Contextualizing. For other stakeholders (e.g. medical care workers, families), ECR data were also insufficient in meeting their respective care needs. For families, their main goals were learning seniors' condition remotely and providing necessary helps and care when needed. Yet, existing ECR data were hard to support these needs due to the low accessibility and accountability of ECR data. That is, there were obvious gaps between ECR data and what families needed, mainly reflecting in the following three points.

First, ECR data were too detailed and professional, which made them not very accessible for families. Specifically, too many detailed data presented in ECR made many families feel hard to identify the meaningful contents, and the relatively professionalism of data further aggravated the low accessible issue of ECR data because most families were without professional background and hard to understand those data very well. F7, whose mother lived in izhaohu due to Parkinson, explained that, "*we are not doctor, can't figure out very well what the data means.*" Second, formal data in ECR usually lacked the necessary situated and contextual information which made them unaccountable for families. As F7 said, "*those data* (in ECR) *are often not clear. It provides data, but doesn't provide why. After I see them, I still don't know what happened and what should I do"*. Third, many families in our study described that the personalized data they needed were absent in ECR, such as "*I always know my mom's hypertension and don't need these data tell me again. My main concern is her blood sugar but it is not in the routine data*" [F5], and "*My mom's condition is relatively good, although she is suffering from Alzheimer's. My only concern is whether she is happy today* [F4]." Due to these reasons, most of our family participants told us they didn't use izhaohu's ECR app very frequently. Some of them even hardly opened it.

For medical care providers, the situation was very similar. Doctors, as the professional, were mainly in charge of checking in the clinic conditions of seniors, and generating professional and precise care programs and supports. The doctors in our study also expressed the data in ECR didn't function very well in supporting them to learn seniors' conditions remotely and then generate the precise care programs. M1, a doctor, considered data-driven precision healthcare were very hard, because "*the current data were too general and lack of fine-grained data, which made them far away from the term of 'precision'*" [M1]. She used '*mobility difficulty*' to explain:

> "*Medically speaking, 'mobility difficulty' includes varying degrees, such as 'can stand, but cannot walk alone', 'can walk two or three feet, or ten feet, or can keep walking for a long time', 'can walk, but cannot go to toilet alone' etc. These are pivotal to precision healthcare. Based on these, we will know what services and assistive technologies should be provided. But in ECR, it was usually one term – 'mobility difficulty'.*" [M1]

Under such circumstances, frontline care workers then performed a series of informal data practices, such as data interpreting, contextualizing, and supplementing through phone calling, communicating in-person, sending WeChat message, etc., to meet stakeholders' needs. WeChat mediated communication, for instance, was the main informal data exchanging way between frontline care workers and families. Most families had added caregivers' personal WeChat account for keeping communication with them. The communicating contents spanned from care-related questions such as "*why did my mom go to toilet so much times last night*" "*My mother's blood pressure seems like a little bit high. How does she feel?*" to work-related questions "*why is our cost this month higher than last month?*" etc.

Our findings suggested WeChat mediated informal data exchanging and communicating was actually the main data sharing and communicating channel in izhaohu. Its functions and roles in meeting stakeholders' data needs were far beyond the official ECR app. Further, we found WeChat mediated informal exchanging and communicating had been treated as an official way. In izhaohu, an official WeChat group (see Fig. 4-c), composed of all frontline care workers, doctors, care managers and families, had been build and used for a long time to support data sharing and communicating among stakeholders. Every day, frontline care workers posted seniors' daily activities (e.g. eating, activities, etc.) in the chatting room in forms of photos, videos, or texts, and answered stakeholders' questions (e.g. families' "*why did my mom seem like not very happy?*" "*Why didn't Grandpa L take part in the morning exercises?*", and doctors' "*please providing seniors' medication data last week*" etc.).

The stakeholders (e.g. families) highlighted the significant roles of these informal data sharing and exchanging in meeting their needs of caring remotely, because they were more convenient, direct, and reliable for them to obtain more rich and accessible data, and ask their questions directly. Yet, for frontline care workers, this ubiquitous data work was "*too heavy*" and increasingly encroached their personal space and time. They needed to deal with various informal data work, e.g. respond families' message, any time any where. In our interviews, most of care workers used "*no personal life*" to described their current work and life.

In summary, for both frontline care workers themselves and stakeholders (e.g. families and doctors), ECR data didn't function very well in supporting them reach their care goals. We identified a series of breakdowns when using data technologies such as ECR to drive care-taking practices and meet stakeholders' goals, and various informal data practices performed by frontline care workers to repair these breakdowns. These informal data work played crucial roles in supporting stakeholders' needs and making data-driven LTC work on the one hand, but increased frontline care workers' workload greatly on the other hand.

6 DISCUSSION

This paper examined the ongoing trend of data-driven healthcare from frontline care workers' perspectives. We reported the gaps between the promise and reality of data-driven LTC, how frontline care workers, as the actual crux of a data-driven configuration, bridge these gaps through a tremendous amount of data work, and how this data work effects their care work. We now provide an in-depth discussion about the practices, problems, and opportunities of care workers' data work in LTC.

6.1 Data Work in Data-driven Long-Term Care: Practices

This paper has demonstrated that the trend of data-drivenness complicates the professional identities of frontline care workers by adding another dimension: data worker. With this identity, a significant amount of data work [6] is conducted to address the tensions between the promised data-driven healthcare vision and the practical care work at different levels and scales, and to underpin and implement data-driven collaborative healthcare. Data work here, we argue, goes beyond the "clerical work" [47] and functions also as a labor of maintenance, articulation, and repair, that both guarantees data technologies' functionalities and acts as an interface between stakeholders.

First, data work constructs and maintains data infrastructures in LTC. Because much data required by LTC are hard to be obtained by sensing technologies directly, frontline care worker then conduct a series of data work to manually produce, input, mend and update these data, including both care-related data (e.g. seniors' health and ADLs) and work-related data (e.g. their working hour and performance). This data work provides the essential "oil" for data-driven healthcare vision. Second, data work functions as an articulation of stakeholders. The process of LTC depends on the collaborations of frontline care workers, doctors, families and others. As the only roles at the scene, frontline care workers then perform data work such as data communicating, sharing, exchanging, etc., to gather and mesh multiple demands and opinions, to bring their distributed efforts together, and to make the data-driven multiple collaboration work [64]. Third, data work functions as a labor of repair. Due to the incorrect and incomplete data, and the lack of valuable, personalized and contextualized contents, the data in ECR system are insufficient for supporting stakeholders to reach their own care goals. Frontline care workers then conduct a lot of efforts to seek, exchange, interpret and contextualize ECR data, and make data meaningful and useful to support collaborative work.

In form, frontline care worker's data work includes both formal and informal parts. Formal data work refers to those the company officially requires care workers do, such as recording daily health, wellbeing and behavioral data every day. It is the part of care workers' work and associates with their work KPI directly. Informal data work refers to those situated and temporary data practices. It doesn't have clear time and space boundaries, and usually be triggered by the breakdowns of data infrastructures. In other words, it might occur whenever and wherever and continues through the whole data-driven healthcare process. Even some informal data work is slowly become formal. For instance, WeChat group mediated data commutation has been treated as formal work, and care workers must share seniors' photos and videos to families through WeChat group every day, and reply the messages in it. Although this work has not yet connected to workers' KPI directly, care workers' performance to it is considered as a reflection of their "work attitude".

Prior literature has examined data work conducted by different roles, including the typical people in healthcare (e.g. doctors, nurse, etc. [6]) and new occupations (e.g. medical secretary, data analyst, data scientist, etc.) [47], in different healthcare contexts (e.g. hospital [47], community healthcare [69], emergency medical services [75] etc.), frontline care workers' data work in LTC is rare in CSCW and HCI. We hope this research provides a first-stepping stone in this direction. The perspective of data work also provides us a lens to examine the effectiveness of data-driven healthcare and to explain the increasing reported cumbersome usability of data technologies in healthcare. Our findings suggest the future design of healthcare technologies should focus on not only how to support care workers' own work, but also the new/extra work generated by using data technologies.

6.2 Data Work in Data-driven Long-Term Care: Problems

While a growing body of literature has paid attention on the detailed practices of data work in different data-driven settings (e.g. [63, 69, 75]), and its significant role in making data useful and meaningful and making 'data-drivenness' become real [2, 6, 40], what we want to highlight in this study is the impacts, especially the negative and downstream effects to performers and to the sustainable development of data-driven healthcare ecosystem.

In the context of LTC, it is frontline care workers – usually female, middle-aged and less-educated migrant labor [1, 53], carrying out this significant data work. For them, besides the impacts what prior literature has reported, such as increased burden [69], we argue a form of labor exploitation is happening to them and being normalized in name of data innovation [42]. To be specific, compared to its "enormous contribution", this data work frontline care worker performed is often invisible [13], undervalued, unpaid, and unacknowledged by healthcare companies and managers [47]. It is considered as frontline care workers' default work. They got neither the necessary resources nor the proper compensation from this labor. They carry out data work not because they want to make data-driven healthcare work, but they are required by the managers. If they don't perform this data work well, e.g. they don't record seniors' daily care data into system carefully, or don't reply families' message in time, they will be considered as "taking a bad attitude to work", be punished and even be quitted. This generates the greater social unfairness issue. The consequence of this unfairness is serious especially under the social statues of the global caregiver shortage [7, 11].

Further, we also argue that if we continue to develop data-driven healthcare ecosystem through this model, without understanding or addressing care workers' massive data work, the future of data-driven LTC will be impoverish. As what we shown in our study, the ubiquitous data work made most frontline care workers feel exhausted. They had to finish these "tasks" in their rest or off-duty time. Their personal lives were also encroached by various informal data work. Meanwhile, almost no incentive mechanism, no matter financial or psychological, was used to incent their data work. They also couldn't experience these promised benefits very well, and therefore preferred to take themselves as 'data workers' [28], instead of 'data beneficiary'. These – increased workload, increased unfairness, the potential labor exploitation, lack of incentives and the limited perceptions to data values – cause negative feelings and attitudes of frontline care workers to data and data work. These negative attitudes would directly impact their motivation to explore and contribute data, and the quality of their data work, which, as a kind of what Sambasivan et al. defined as "data cascades" [62], would then impact the accuracy and efficiency of data-driven healthcare.

This negative loop is very harmful because frontline care workers are the actual ones who know the richest data information about seniors in LTC, and their data practices would directly impact the quality, transparency, accountability of data [63]. We have identified some data cascades [62] in our study. For instance, because frontline care workers considered data capturing and recording as "tasks" related to their KPI with the above-mentioned negative attitudes, they performed data work with a task-finishing mindset, instead of providing the rich and valuable data proactively. They might ignore the detailed and context data because recording these data meant more workload. They even improvised some contents or concealed some important materials (e.g. seniors' accident) because that might cause "bad data" and impact their KPI. These would cause various data issues, such as introducing unpredictable noises into the dataset, increasing data sparsity, and reducing data accountability and transparency. All these data issues would cause negative and downstream effects to data-driven healthcare [62].

As data-drivenness are being increasingly promised, more people would be involved into this data work. We strongly suggest considering data-driven vision beyond quality, efficiency, effectiveness, etc. to a broader data-driven ecosystem, and paying attention to the invisible data work in data-drivenness and its impacts to performers, especially to "low-level" workers. By stressing data work as labor, we hope to render data work visible and acknowledge, which we believe is crucial to enable more just, ethical and useful data future.

6.3 Data Work in Data-driven Long-Term Care: Opportunities

The nature of data work is a kind of human engaged articulation work [68]. Through the lens of data work, we identify a series of breakdowns of data and data technologies in izhaohu's ECR system, mainly reflecting in the tensions between formal data and practical needed data, and between data-driven configuration and practical care process. We consider these breakdowns as the main reasons of care workers' burdensome data work. Besides the problematic or imperfect data and data system, we think the deeper reasons might be the inappropriate relationships between people, practices, and data. Unpacking these issues provides us opportunities to reconsider these relationships and the further direction of data-driven healthcare.

Quality vs. Quantity Aspect of Data. Given the significant promises of datafication upon healthcare, a large body of work has turned to gather and store "big data", and then uses data drive healthcare and management. Yet, in the real-world LTC practices, most our participants did not really consider how quantity of healthcare data would matter. More fundamentally, they seemed to have a vague sense of what quantity meant in their caregiving work. Some even expressed huge amount of data posted a distinct challenge to the actionability [20] of data: when they had some key data, they could judge the situation, predict the risk or intervene precisely and timely; But, when they had huge amount of data, they didn't know what that meant. In comparation, almost all our participants reported their high demands to quality, meaningful, and accountable data. For example, they all considered seniors' personal, socio-psychological, emotional, and contextual data were more valuable in supporting their day-to-day care work. Meanwhile, our findings also showed the expectations, experiences and perceptions to data contents and granularity were vary with each role.

In the view of "data activism", the more data, the better performance and values. The big-data fundamentalists even declare through properly mining, large data sets could recognize correlations and patterns, predict behaviors and results, address long-standing challenges in healthcare, and even replace the need for theory and science [4]. Consequently, building 'big data sets' has become a critical issue in contemporary data-driven society [60]. Yet, our findings, in line with recent literature (e.g. [2, 72]), provide the evidences that while human and material resources are consumed to promote datafication, its functions and values for practical care work are limited by the quality of data. Considering this, as well as the possible data privacy and ethical issues, our study suggests a careful consideration about the reconceptualization of data paradigm in data-driven healthcare, especially the boundary of quantified data in LTC.

Caregiving Practices Driving vs. Driven by Data. From the perspective of data power [60], we identity the asymmetric relations [14, 15] between those who collect and store data (frontline care workers) and those whom data collection target (families, managers and professional workers), as well as the divide between 'data rich' (e.g. managers, professional workers and enterprises) and 'data poor' (individuals). Specifically, frontline care workers are the practical 'data rich', but the technological 'data poor'. What happened then is an ambivalent data-driven healthcare process: they are asked to transfer their rich care data into system (data work) under the pre-defined framings, then be "guided" by these data about how to care. This process is less effective and brings a series of problems.

As data-driven healthcare is growing rapidly, people, data and practices will be further intertwined. Our findings, as well as the challenges of data-driven healthcare technologies identified in prior research [72, 77] raise an urgent need to re-consider the relationship between human, data and technologies in healthcare, especially the roles of data and data technologies in LTC. This issue is important to not only care workers' workplace wellbeing, but also the healthy and sustainable data-driven healthcare future.

6.4 Limitations and Future Work

This paper presents a preliminary investigation of frontline care workers' data work in the real-world data-driven LTC setting. Three limitations should be noted. First, our study was conducted in China. While the specific socio-political context in China provided us a unique chance to gain more insights into data-driven LTC, it, as well as some specific features of izhaohu (e.g. organizational structure and management philosophy) might influence the generalization of data work we studied. Nevertheless, we believe as a popular data-driven LTC mode in China, our study provides HCI and CSCW communities a valuable case for improving the further development of datafication in LTC. Second, the three-week observation study might cause the limited exposure, for instance, to different times of the year (e.g. a holiday or the seasons), and to more diverse care events. Third, our examination to frontline care workers' data work mainly focused on their physical practices and lacked attention to their socio-emotional and psychological factors, which were actually crucial to frontline care workers' work and life. In the future, we will examine data and data technologies in multiple data-driven long-term care ecosystems to generate a more comprehensive picture of data-driven LTC. We will also future examine frontline care workers' data work from the psychosocial perspective to deepen our understanding to their data work.

7 CONCLUSION

In this study, we conducted a qualitative study to examine the ongoing data-driven healthcare in a real-world LTC facility in China. While data-driven technologies are well-meaning and promising "better" healthcare – improving cost-effective, quality, efficiency, etc., our study demonstrates care workers, especially frontline care workers, are actually the crux of this data-driven configuration. For implementing 'data-driven vision' in practice, they (are required to) conduct a series of data work. We uncover that under the promised data-driven future, a kind of implicit labor exploitation is happening and normalized in the name of data innovation, which should be seriously taken by not only researchers in CSCW and related fields, but also organizations and people with vested interests in data and data technologies. We discuss the relationship between people, data and practices in the context of long-term care setting and design implication for maintaining the sustainable data-driven LTC.

ACKNOWLEDGMENTS

We deeply appreciate izhaohu Co., Ltd. for the support to the overall study and care workers in it for their time. We are grateful for the valuable feedback and comments provided by the anonymous reviewers. This is work is supported by Shanghai Science and Technology Innovation Action Plan Project (21511104500).

REFERENCES

[1] AARP. 2015. Caregiving in the U.S. https://www.aarp.org/content/dam/aarp/ppi/2015/caregiving-in-the-us-research-report-2015.pdf.

[2] Adegboyega Akinsiku, Ignacio Avellino, Yasmin Graham and Helenam Mentis. 2021. It's not just the movement: experiential information needed for stroke telerehabilitation. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI 2021), 1-12. DOI: https://doi.org/10.1145/3411764.3445663.

[3] Jessicas Ancker, Yolanda Barron, Maxinel Rockoff, Diane Hauser and Michelle Pichardo, et al. 2011. Use of an electronic patient portal among disadvantaged populations. JOURNAL OF GENERAL INTERNAL MEDICINE 26, 10 (2011), 1117-1123.

[4] Chris Anderson. 2008. The end of theory: The data deluge makes the scientific method obsolete. Wired magazine 16, 7 (2008), 16-17.

[5] Javier Andreu-Perez, Carmency Poon, Robertd Merrifield, Stephentc Wong and Guang-zhong Yang. 2015. Big data for health. IEEE Journal of Biomedical and Health Informatics 19, 4 (2015), 1193-1208.

[6] Claus Bossen, Kathleenh Pine, Federico Cabitza, Gunnar Ellingsen and Enricomaria Piras. 2019. Data work in healthcare: An Introduction. SAGE Publications Sage UK: London, England, 465-474. DOI: https://doi.org/10.1177/1460458219864730

[7] Bailey Bryant. 2019. Caregiver Shortage Could Mean 7.8 Million Unfilled Jobs By 2026. https://homehealth-carenews.com/2019/01/caregiver-shortage-could-mean-7-8-million-unfilled-jobs-by-2026/

[8] Asa Cajander and Christiane Grunloh. 2019. Electronic health records are more than a work tool: conflicting needs of direct and indirect stakeholders. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI 2019), 1-13. DOI: https://doi.org/10.1145/3290605.3300865

[9] Neenal Chappell and Rcolin Reid. 2002. Burden and well-being among caregivers: examining the distinction. The Gerontologist 42, 6 (2002), 772-780. DOI: https://doi.org/10.1093/geront/42.6.772

[10] Yunan Chen, Victor Ngo and Sunyoung Park. 2013. Caring for caregivers: designing for integrality. In Proceedings of the 2013 conference on Computer supported cooperative work (CSCW 2013), 91-102. DOI: https://doi.org/10.1145/2441776.2441789

[11] Cassiem Chew. 2017. Caregiver shortage reaches critical stage. Provider 43 (2017), 14-28.

[12] General Office of China. 2021. The 14th Five-Year Plan and the Vision 2035 of China. https://www.ndrc.gov.cn/xxgk/zcfb/ghwb/202103/t20210323_1270124.html?code=&state=123.

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

[14] Kate Crawford, Maryl Gray and Kate Miltner. 2014. Critiquing Big Data: Politics, ethics, epistemology| special section introduction. International Journal of Communication 8 (2014), 10.

[15] Jose Van Dijck. 2014. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. Surveillance & society 12, 2 (2014), 197-208. DOI: https://doi.org/10.24908/ss.v12i2.4776

[16] Pamela Doty, Korbin Liu and Joshua Wiener. 1985. An overview of long-term care. Health Care Financing Review (Health Care Financ Rev) 6, 3 (1985), 69.

[17] Paula England. 2005. Emerging theories of care work. Annual Review of Sociology (Annu. Rev. Sociol.) 31 (2005), 381-399.

[18] Lynn Etters, Debbie Goodall and Barbarae Harrison. 2008. Caregiver burden among dementia patient caregivers: a review of the literature. Journal of the American Academy of Nurse Practitioners 20, 8 (2008), 423-428. DOI: https://doi.org/10.1111/j.1745-7599.2008.00342.x

[19] Sanford Finkel, Saraj Czaja, Zoran Martinovich, Carol Harris and Donna Pezzuto, et al. 2007. E-care: a telecommunications technology intervention for family caregivers of dementia patients. The American journal of geriatric psychiatry 15, 5 (2007), 443-448. DOI: https://doi.org/10.1097/JGP.0b013e3180437d87

[20] Brittany Fiore-Gartland and Gina Neff. 2015. Communication, mediation, and the expectations of data: Data valences across health and wellness communities. International Journal of Communication 9 (2015), 19.

[21] Pinsym Foong, Shengdong Zhao, Felicia Tan and Josephjay Williams. 2018. Harvesting caregiving knowledge: Design considerations for integrating volunteer input in dementia care. In Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI 2018), 1-12, DOI: https://doi.org/10.1145/3173574.3173653

[22] Roland Gamache, Hadi Kharrazi and Jonathanp Weiner. 2018. Public and population health informatics: the bridging of big data to benefit communities. Yearbook of medical informatics 27, 01 (2018), 199-206.

[23] Barney Glaser (Ed.). 1978. Theoretical sensitivity: advances in the methodology of grounded theory. Sage Publications, CA, USA.

[24] Oliver Gruebner, Martin Sykora, Sarahr Lowe, Ketan Shankardass and Sandro Galea, et al. 2017. Big data opportunities for social behavioral and mental health research. SOCIAL SCIENCE & MEDICINE (2017).

[25] Gaspard Harerimana, Beakcheol Jang, Jongwook Kim and Hungkook Park. 2018. Health big data analytics: A technology survey. IEEE Access 6 (2018), 65661-65678.

[26] Gillianr Hayes, Kareng Cheng, Senh Hirano, Karenp Tang and Marnis Nagel, et al. 2014. Estrellita: a mobile capture and access tool for the support of preterm infants and their caregivers. ACM Transactions on Computer-Human Interaction (TOCHI 2014) 21, 3 (2014), 1-28. DOI: https://doi.org/10.1145/2617574

[27] Christian Heath and Paul Luff. 1996. Documents and professional practice: “bad” organisational reasons for “good” clinical records. In Proceedings of the ACM conference on Computer supported cooperative work (CSCW 1996), 354-363, DOI: https://doi.org/10.1145/240080.240342

[28] Klaus Hoeyer and Sarah Wadmann. 2020. 'Meaningless work': How the datafication of health reconfigures knowledge about work and erodes professional judgement. ECONOMY AND SOCIETY 49, 3 (2020), 433-454. DOI: https://doi.org/10.1080/03085147.2020.1733842

[29] Torhild Holthe, Rita Jentoft, Cathrine Arntzen and Kirsten Thorsen. 2018. Benefits and burdens: family caregivers ’ experiences of assistive technology (AT) in everyday life with persons with young-onset dementia (YOD). Disability and Rehabilitation: Assistive Technology 13, 8 (2018), 754-762. DOI: https://doi.org/10.1080/17483107.2017.1373151

[30] Taranu Ionut. 2016. Data mining in healthcare: decision making and precision. Database Systems Journal 6, 4 (2016), 33-40.

[31] Azra Ismail and Neha Kumar. 2018. Engaging solidarity in data collection practices for community health. Proceedings of the ACM on Human-Computer Interaction 2, CSCW (2018), 1-24. DOI: https://doi.org/10.1145/3274345

[32] Andersfogh Jensen and Dennis Normark (Ed.). 2018. Pseudoarbejde: Hvordan vi fik travlt med at lave ingenting. Gyldendal Business.

[33] Ashishk Jha, Catherinem DesRoches, Ericg Campbell, et al. 2009. Use of electronic health records in US hospitals. New England Journal of Medicine (NEJMAG) 360, 16 (2009), 1628-1638.

[34] Elizabeth Kaziunas, Marks Ackerman, Silvia Lindtner and Joycem Lee. 2017. Caring through data: Attending to the social and emotional experiences of health datafication. In Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW 2017), 2260-2272. DOI: https://doi.org/10.1145/2998181.2998303

[35] Elizabeth Kaziunas, Michael S. Klinkman, and Mark S. Ackerman. 2019. Precarious Interventions: Designing for Ecologies of Care. In Proceedings of the ACM on Human-Computer Interaction, Vol. 3, No. CSCW, Article 113, 2019, 1-27. DOI: https://doi.org/10.1145/3359215

[36] Predrag Klasnja and Wanda Pratt. 2012. Healthcare in the pocket: mapping the space of mobile-phone health interventions. JOURNAL OF BIOMEDICAL INFORMATICS 45, 1 (2012), 184-198. DOI: https://doi.org/10.1016/j.jbi.2011.08.017

[37] Jacqueline LaPointe. 2020. Long-Term Care Demand to Double Despite Workforce Constraints. https://revcycleintelligence.com/news/long-term-care-demand-to-double-despite-workforce-constraints.

[38] Florian Ldenpfennig, Francisco Nunes and Geraldine Fitzpatrick. 2015. ProxyCare: Integrating informal care into formal settings. In 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth 2015), IEEE, 141-144. DOI: https://doi.org/10.4108/icst.pervasivehealth.2015.259109

[39] Garyr Lee and Masako Ishii-Kuntz. 1987. Social interaction, loneliness, and emotional well-being among the elderly. RESEARCH ON AGING 9, 4 (1987), 459-482. DOI: https://doi.org/10.1177/0164027587094001

[40] Ashlinj Lee and Petas Cook. 2020. The myth of the "data-driven" society: Exploring the interactions of data interfaces, circulations, and abstractions. Sociology Compass 14, 1 (2020), e12749. DOI: https://doi.org/10.1111/soc4.12749

[41] Znaonui Liang, Gang Zhang, Jimmyxiangji Huang and Qmmingvivian Hu. 2014. Deep learning for healthcare decision making with EMRs. In Proceedings of 2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 556-559. DOI: https://doi.org/10.1109/BIBM.2014.6999219

[42] Silvia Lindtner. (Ed.). 2020. Seeing Like a Peer: Happiness Labor and the Microworld of Innovation. Prototype Nation, Princeton University Press, 144-171.

[43] Alexjiahong Lu, Tawannar Dillahunt, Gabriela Marcu and Marks Ackerman. 2021. Data work in education: Enacting and negotiating care and control in teachers' use of data-driven classroom surveillance technology. Proceedings of the ACM on Human-Computer Interaction 5, CSCW2 (2021), 1-26. DOI: https://doi.org/10.1145/3479596

[44] Deborah Lupton. 2016. The diverse domains of quantified selves: self-tracking modes and dataveillance. ECONOMY AND SOCIETY 45, 1 (2016), 101-122. DOI: http://dx.doi.org/10.1080/03085147.2016.1143726

[45] Milagros Miceli, Julian Posada and Tianling Yang. 2022. Studying up machine learning data: Why talk about bias when we mean power? Proceedings of the ACM on Human-Computer Interaction 6, GROUP (2022), 1-14. DOI: https://doi.org/10.1145/3492853

[46] Briand Modena, Otmane Bellahsen, Nima Nikzad, Angela Chieh and Nathan Parikh, et al. 2018. Advanced and accurate mobile health tracking devices record new cardiac vital signs. HYPERTENSION 72, 2 (2018), 503-510. DOI: https://doi.org/10.1161/HYPERTENSIONAHA.118.11177

[47] Naja Moller and Trine Nielsen. 2020. Who Does the Work of Data? Health Matters (2020).

[48] Wben Mortenson, Alex Pysklywec, Marcusj Fuhrer, Jeffreyw Jutai and Michelle Plante, et al. 2018. Caregivers' experiences with the selection and use of assistive technology. Disability and Rehabilitation: Assistive Technology 13, 6 (2018), 562-567. DOI: https://doi.org/10.1080/17483107.2017.1353652

[49] Lifetimes News. 2017. China's pension pressure is fifth in the world. http://cpc.people.com.cn/n1/2017/0221/c64387-29095133.html.

[50] Andrewb Neang, Will Sutherland, Michaelw Beach and Charlottep Lee. 2021. Data Integration as Coordination: The Articulation of Data Work in an Ocean Science Collaboration. Proceedings of the ACM on Human-Computer Interaction 4, CSCW3 (2021), 1-25. DOI: https://doi.org/10.1145/3432955

[51] Renmin News. 2021. General Office of the State Council of China. Building Full-coverage, High-quality, Professional, and Sustainable Ageing Care Facilities to Cope with Ageing Boom. http://www.gov.cn/xinwen/2021-04/01/content_5597224.htm.

[52] Renmin News. 2020. Technologies Supported Chinese Ageing Care. http://www.stdaily.com/cehua/qzcyj/2020-10/20/content_1030006.shtml.

[53] Xinhua News. 2020. We need more caregivers. http://www.xinhuanet.com/politics/2020-09/03/c_1126446000.htm.

[54] Renmin News. 2021. Building Full-coverage, High-quality, Professional, and Sustainable Ageing Care Facilities to Cope with Ageing Boom. http://www.gov.cn/xinwen/2021-04/01/content_5597224.htm.

[55] Liuqian Ni, Yuling Sun, Yanqin Yang and Liang He. 2019. Human-Engaged Health Care Services Recommendation for Aging and Long-term Care. In Proceedings of IEEE 23rd International Conference on Computer Supported Cooperative Work in Design, 339-344. DOI: https://doi.org/10.1109/CSCWD.2019.8791885

[56] Edwardc Norton. 2000. Long-term care. Handbook of health economics 1 (2000), 955-994.

[57] World Health Organization. 2015. China country assessment report on ageing and health. https://apps.who.int/iris/bitstream/handle/10665/194271/9789241509312_eng.pdf.

[58] World Health Organization. Noncommunicable diseases. https://www.who.int/health-topics/noncommunicable-diseases#tab=tab_1.

[59] Jeffery Reeves, Hannahm Hollandsworth, Francescaj Torriani, Randy Taplitz and Shira Abeles, et al. 2020. Rapid response to COVID-19: health informatics support for outbreak management in an academic health system. JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION 27, 6 (2020), 853-859. DOI: https://doi.org/10.1093/jamia/ocaa037

[60] Minna Ruckenstein and Natashadow Schull. 2017. The datafication of health. Annual Review of Anthropology (ARAPCW) 46 (2017), 261-278. DOI: https://doi.org/10.1146/annurev-anthro-102116-041244

[61] Annadisgreta Rudolfsdottir. 2000. 'I am not a patient, and I am not a child': the institutionalization and experience of pregnancy. FEMINISM & PSYCHOLOGY 10, 3 (2000), 337-350.

[62] Nithya Sambasivan, Shivani Kapania, Hannah Highfill, et al. 2021. “Everyone wants to do the model work, not the data work”: Data Cascades in High-Stakes AI. In proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI 2021), 1-15. DOI: https://doi.org/10.1145/3411764.3445518

[63] Morganklaus Scheuerman, Alex Hanna and Emily Denton. 2021. Do datasets have politics? Disciplinary values in computer vision dataset development. Proceedings of the ACM on Human-Computer Interaction 5, CSCW2 (2021), 1-37. DOI: https://doi.org/10.1145/3476058

[64] K Schmidt and L Bannon. 1992. Taking CSCW seriously. Computer Supported Cooperative Work (CSCW). An International Journal. 1, 1--2 (1992), 7-40.

[65] Tamar Sharon. 2017. Self-tracking for health and the quantified self: Re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. Philosophy & Technology 30, 1 (2017), 93-121.

[66] Ida Sim. 2019. Mobile devices and health. New England Journal of Medicine (NEJMAG) 381, 10 (2019), 956-968.

[67] Stephen Kaye, Charlene Harrington and Mitchell LaPlante. 2010. Long-term care: Who gets it, who provides it, who pays, and how much? HEALTH AFFAIRS 29, 1 (2010), 11-21. DOI: https://doi.org/10.1377/hlthaff.2009.0535

[68] Lucy Suchman. 1996. Supporting articulation work.

[69] Nervo Verdezoto, Naveen Bagalkot, Syedazainab Akbar, Swati Sharma and Nicola Mackintosh, et al. 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 (2021), 1-31. DOI: https://doi.org/10.1145/3449165

[70] Gemma Webster and Vickil Hanson. 2014. Technology for supporting care staff in residential homes. ACM Transactions on Accessible Computing (TACCESS) 5, 3 (2014), 1-23. DOI: https://doi.org/10.1145/2543577

[71] Jamesb Wetmore, Suying Li, Heng Yan, Muna Irfan and Nazia Rashid, et al. 2019. Increases in institutionalization, healthcare resource utilization, and mortality risk associated with Parkinson disease psychosis: Retrospective cohort study. PARKINSONISM & RELATED DISORDERS 68 (2019), 95-101.

[72] Qingyu Xu, Yuling Sun and Liang He. 2021. Design Considerations for Information Collaboration in Long-term Senior Care System. In Proceedings of 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design, IEEE, 1040-1045. DOI: https://doi.org/10.1109/CSCWD49262.2021.9437812

[73] Bong-min Yang. 2009. The future of health technology assessment in healthcare decision making in Asia. PHARMACOECONOMICS 27, 11 (2009), 891-901. DOI: https://doi.org/10.2165/11310280-000000000-00000

[74] Janee Zarem. 2010. Today's continuing care retirement community (CCRC). CCRC Task Force, American Seniors Housing Association (2010).

[75] Zhan Zhang, Karen Joy, Pradeepti Upadhyayula, Mustafa Ozkaynak and Richard Harris, et al. 2021. Data Work and Decision Making in Emergency Medical Services: A Distributed Cognition Perspective. Proceedings of the ACM on Human-Computer Interaction 5, CSCW2 (2021), 1-32. DOI: https://doi.org/10.1145/3479500

[76] Jie Zhou, Yuling Sun and Liang He. 2018. A Data Driven Collaborative Caring Framework for Aging in Place in China. In Proceedings of 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design, IEEE, 359-364. DOI: https://doi.org/10.1109/CSCWD.2018.8465191

[77] Xiaomu Zhou, Marks Ackerman and Kai Zheng. 2010. Doctors and psychosocial information: records and reuse in inpatient care. In proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2010), 1767-1776. DOI: https://doi.org/10.1145/1753326.1753592

Received January 2022; revised July 2022; accepted November 2022.