



The disciplined customer: A video-based study of automated self-service hotels

new media & society
2024, Vol. 26(9) 5013–5038
© The Author(s) 2024



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14614448241251793
journals.sagepub.com/home/nms



Xinzhi Xu*

The Chinese University of Hong Kong, Hong Kong

Christian Greiffenhagen* 

The Hong Kong Polytechnic University, Hong Kong

Abstract

This article studies customers entering automated self-service hotels in China and using a facial recognition kiosk for registration. Based on video recordings of 674 cases of customers checking in, we show that, as is common in self-service, customers need to do work that was previously done by hotel staff: They are working customers. We then argue that, when interacting with the facial recognition kiosk, customers are also doing something more: First, they *present* themselves to the machine by, for example adjusting their standing position or appearance; second, they *perform* for the machine by following its instructions, closing their eyes or opening their mouths; finally, they *express* various emotions towards the machine, such as anger or embarrassment. In sum, we show that customers are working not just *with* their bodies but also *on* their bodies, which are being disciplined by the machine. They become ‘disciplined customers’.

Keywords

Automation, discipline, facial recognition, human–computer interaction, human–machine communication

*These authors have contributed equally to this work and are listed in reverse alphabetical order.

Xinzhi Xu is also affiliated to Jinan University, Canton, China

Corresponding author:

Christian Greiffenhagen, Department of Applied Social Sciences, The Hong Kong Polytechnic University, Kowloon, Hong Kong.

Email: christian.greiffenhagen@polyu.edu.hk

Introduction

Recent decades have seen the widespread use of automated communication technologies: from voice assistants (e.g. Amazon's Alexa and Apple's Siri) to chatbots used for online customer service and social robots such as Kismet or Bandit. These technologies are designed to 'communicate' with humans, to 'understand' and 'recognise' them, in order to 'serve the needs of human communication' (Hepp, 2020: 1411). In other words, these systems are 'communicative subjects, instead of mere interactive objects' (Guzman and Lewis, 2020: 71). These technologies are part of an increasing automation of all aspects of social life (Andrejevic, 2019). While automation may have started in the manufacturing industry, it has now also entered the service industry (Smith, 2020) as well as the home (Brush et al., 2011).

There are many fantastic tales told about automation, in particular, that automation will completely replace human workers in many sectors. However, as Munn (2022) powerfully argues, most of the discourse about automation is a myth. There is a huge mismatch between the aspirations of automation technologies and the realities on the ground (Gray and Suri, 2019: xxii). Automation most of the time is not fully 'automatic' but relies on a lot of human labour behind the scenes to make the automation work, 'to smooth out technology's rough edges' (Mateescu and Elish, 2019: 5). For example, in order for robots to be able to take care of the elderly, people need first to take a lot of care of the robots (Lipp, 2023). However, such human labour is typically invisible and hidden (Altenried, 2022: 119; Munn, 2022: 4), which is why Gray and Suri (2019) refer to it as 'ghost work'.

Automation typically does not replace human work, but instead *reconfigures* it in various ways (Altenried, 2022: 6; Gray and Suri, 2019: xxii; Mateescu and Elish, 2019: 5; Munn, 2022: 33; Smith, 2020: 13). Most research has focussed on how automation reconfigures and changes what workers do. Less studied is how automation changes what consumers do. This has been explored by researchers in the areas of human-computer interaction (HCI) and human-machine communication (HMC) who investigate how users interact with automated communication technologies such as Microsoft Kinect (Harper and Mentis, 2013), Amazon's Alexa (Porcheron et al., 2018) or Aldebaran's robot *Nao* (Pelikan and Broth, 2016) naturally in everyday settings. Those studies have focussed on the 'work' done by the human users in these interactions.

In this article, we study what customers need to do in order to check into a fully automated self-service hotel in China, which uses a facial recognition kiosk to verify their identity. While facial recognition is often used to identify unknown people and is therefore associated with (state) surveillance, it can also be used for the purpose of verifying that people are who they claim to be (Gates, 2011: 18), for instance in cases of the 'face unlock' function on mobile phones and access control for certain rooms and buildings. As Bucher (2022) puts it: 'While facial recognition technology is often linked to the realm of state control, contemporary consumer technology shows how much it has become an integral part of our everyday lives' (p. 644). We explore how customers interact with facial recognition kiosks to check into the hotel and discuss how this interaction changes the role of the customer. We do so in the context of China, where facial recognition is extremely common. China is not only one of the countries with the most



Figure 1. A customer using the facial recognition kiosk.

closed-circuit television (CCTV) cameras (Kostka et al., 2021) but also a country where facial recognition is used for many everyday activities, including getting access to buildings (He et al., 2019) and payment in stores (Hu et al., 2023). Furthermore, Chinese citizens express fewer concerns towards the use of such technologies than citizens in other countries (Kostka et al., 2021; Steinhardt et al., 2022).

Our study is based on qualitative video analysis (Heath et al., 2010). We collected over 70 hours of video data with 2 GoPro cameras, recording 674 cases of customers checking in. Figure 1 provides an example.

Akin to using a passport photo booth, customers start by preparing themselves in various ways, which includes adjusting their standing position (so they can be captured by the camera) as well as working on their appearance (e.g. by adjusting removing accessories). In other words, customers need to *present* themselves to machine. The facial recognition process requires customers to follow a number of instructions of the machine (such as ‘close eyes’, ‘shake head’ or ‘open mouth’) to be authenticated by the kiosk. Customers are not just presenting themselves to the machine; they are *performing* for the machine. This performance is completely determined by the machine, which tells customers what to do, when and how often. When customers are not immediately recognised, they need to modify their bodily responses in the hope that their changes will result in a successful recognition (see Greiffenhagen et al., 2023). Ultimately customers’ bodies are being controlled by the machine, which can cause them to feel embarrassed or angry, which in turn can prompt them to express emotional reactions to the machine. Occasionally, they become irate and shout at the kiosk; more frequently, they laugh at their own performance, especially in the presence of bystanders.

Previous research has shown that, when changing from service to self-service, customers are no longer treated as ‘kings’ but instead become ‘working customers’ (Pentzold and Bischof, 2023; Rieder and Voß, 2010), taking up much of the work that was previously done by service personnel. In this article, we argue that automated communication technologies, such as facial recognition, produce customers who are not only ‘working’ but are furthermore ‘disciplined’ by these technologies.

In speaking of ‘discipline’, we want to highlight that customers are doing more than just working *with* their body, as they do in a factory on an assembly line or at a self-service checkout, where they need to use their hands to handle goods in various ways (Andrews, 2019). In our setting, in trying to render themselves ‘understood’ or ‘recognised’ by the facial recognition kiosk, customers alter how they speak, look and even act.

In other words, they are also working *on* their bodies. Their body is the target of what they are being told to do. Their body is being 'reconfigured' (Lipp, 2023), or, as we would put it, 'disciplined', by the system. Our use of discipline is inspired by Foucault's (1978: 139) description of how modern societies have increasingly regulated, controlled and disciplined human bodies in various institutions, from the military to schools. We see our case as an example of such disciplining work of the body, where customers must control their eyes, head and mouth in a way that makes them 'machine-recognisable'.

While Foucault drew attention to the fact that people also discipline themselves even in the absence of a controlling authority, that is become *self-disciplined* (Pylypa, 1998), customers in our setting are disciplined by a technology that uses voice commands to tell them what to do with their bodies. We thus also draw on the more recent literature on 'algorithmic governance' (Katzenbach and Ulbricht, 2019; Ziewitz, 2016), which highlights that our lives are increasingly under the control of algorithms. We are particularly inspired by Aneesh's (2006, 2009) notion of 'algocratic' governance, which he contrasts with 'bureaucratic' and 'panoptic' governance, and which is used to highlight how machines today have authority: 'Unlike the unlettered machines of the industrial age, the new machine has the ability to communicate commands as an authority in addition to faithfully carrying out commands of the worker' (Aneesh, 2006: 110). In today's societies, machines have been endowed with the authority to command, supervise and review people's actions. In our case, the machine has the authority to demand 'appropriate' responses from the customers, who must change what they are doing if the machine treats their response as inadequate.

All technologies, from an automatic teller machine (ATM) to a supermarket self-checkout terminal, require customers to do things in a certain way, that is to conform to the commands of the machine. The facial recognition kiosk in our setting is particularly an explicit form of such discipline. First, since it interacts with the customer through verbal commands, it resembles more the situation where a human tells a customer what to do. Second, since it tells the customer not only to do certain things with their body (e.g. putting a PIN code using a keypad) but also to perform in a certain way in front of a camera (e.g. to shake their head or open their mouth) and thereby to work *on* their body. This article proposes the concept of the disciplined customer to highlight how algorithmic self-service can target individuals' bodies and govern their movements, drawing attention also to the asymmetric relationship between human users and automated machines.

Background

In order to draw out more clearly how our concept of the 'disciplined customer' differs from that of the 'working customer', we review the emergence of self-service and the changing role for employees and customers. The focus of our analysis is a registration kiosk based on facial recognition technology, so we next discuss how this can be used for both surveillance and verification of people's identity. Finally, as discussed in the introduction, while there is a lot of research on how automation reconfigures what workers do, there have been only a few studies that have explored how users interact with these technologies, which we review in the final part of this section.

Self-service and 'working customers'

Self-service was initially mainly used in grocery stores, but it has been extended to many other businesses (Andrews, 2019; Palm, 2016). Today, we encounter self-service in restaurants, gas stations, banks, airports and even hotels (Lukanova and Ilieva, 2019). Self-service is often associated with efficiency, predictability and standardisations; it is thus seen as part of the increasing rationalisation of the service economy, captured in Ritzer's (2013 [1993]) notion of 'McDonaldisation'.

In changing from service to self-service, consumers take up much of the work that was previously done by service employees, typically for free. Customers have become 'working customers' (Rieder and Voß, 2010; see also Dujarier, 2016; Glucksmann, 2016; Goodwin, 1988; Pentzold and Bischof, 2023). Indeed, in modern societies, consumers are responsible for so much that Andrews (2019) speaks of the 'overworked consumer'. Researchers have highlighted that this constitutes a new form of exploitation under capitalism (Ritzer, 2013 [1993]).

There are, however, also various problems associated with the reliance on self-service. Customers often encounter various problems in operating these machines and therefore need assistance from employees. For example, customers using supermarket self-checkout may need help as often as one out of every three times (Andrews, 2019: 107). Furthermore, the introduction of self-checkouts has been found to increase theft and walk-offs (Andrews, 2019: 51), which again has led to new forms of work for employees, for example monitoring of customers to make sure that they do not steal (Mateescu and Elish, 2019: 38).

The standardisation of service stands in direct opposition to the more traditional customisation of service through personal interaction, which can render automation disenchanting for customers. As Andrews (2019) notes: 'efforts to rationalize sales transactions through automation and self-service may be more cost-efficient but frustrating and dehumanising for customers' (p. 51). That is to say, not only do customers 'work', but they also do so in the ways prescribed by the machine, which 'tells the consumer what he must do and how he must do it' (Dujarier, 2016: 561). In that sense, self-service requires what Aneesh (2006, 2009) terms 'algoratic control', wherein customers are controlled by a machine installed by the organisation. Self-service is thus not limited to the exploitation of consumers through unpaid work; it also involves control and discipline. The effect can be disenchanting or dehumanising (Andrews, 2019: 51). As Ritzer (2013 [1993]) says: 'customers might even feel as if they are being fed like livestock on an assembly line' (p. 135). In our analysis of the interaction with the facial recognition kiosk, we focus in particular on how the kiosk disciplines users and possibly thereby disenchanters or dehumanises the hotel visit.

Facial recognition for surveillance and verification

The key technology in self-service hotels is the facial recognition kiosk, which is employed during check-in to identify customers and sometimes to allow them access to their rooms (Lukanova and Ilieva, 2019: 170). Facial recognition technology uses the face 'as an index to identity' (Gates, 2011: 8), treating the face as 'the new fingerprint' (Goriunova, 2019: 127).

Following Gates (2011: 18), we can distinguish between two different applications of facial recognition technologies. The perhaps most known application is for the *identification* of unknown people, which is when facial recognition is used for the purpose of surveillance. In this situation, people may be unaware of it and, more importantly, may not consent. Another use is for *verification*, which involves scanning known people to check whether they are who they claim to be. People are typically aware of this and may want it to succeed in order to speed up what they are trying to do. We encounter this application in passport control (Robertson, 2009), the ‘face unlock’ function on many smartphones (Goriunova, 2019), or the verification of financial transactions (Hu et al., 2023).

Most research on facial recognition has focussed on identification and specifically the case of surveillance (e.g. Bueno, 2020; Gates, 2011; Manokha, 2018). This line of research often draws on Foucault’s writings to highlight the ‘panoptic dimension’ (Bueno, 2020: 78) of facial recognition. Researchers have pointed out that the most pervasive form of surveillance using facial recognition violates people’s rights and privacy. Furthermore, studies have highlighted that automatic facial recognition can often lead to misidentification (Crawford, 2019).

Facial recognition for verification, where people are aware that they are being scanned and indeed actively help that recognition, has received relatively little attention. There has been some research that has explored why facial recognition has become the most prominent form of biometric verification. One reason seems to be that it builds on existing bureaucratic norms and practices (del Rio et al., 2016; Gates, 2011: 46): we are all familiar with a human checking one’s passport against one’s face, which can be done by a machine rather than a human being. Another reason is that facial recognition is considered more convenient for users than other technologies (Kostka et al., 2021). For example, He et al. (2019) investigated people’s experience of using facial recognition in workplaces and found that users appreciated the convenience and efficiency of passing through doors without having to take out additional items for identification.

China is a country where facial recognition is extremely widespread, for both identification and verification. China is one of the countries with the most CCTV cameras in the world, 170 million CCTV cameras were in use in 2018 (Kostka et al., 2021: 674). More importantly for our study, facial recognition is being used in many aspects of everyday life, for example getting access to buildings without using keys or cards (He et al., 2019). Facial recognition can also now be used in many shops, where customers can pay by presenting their face to a screen without having to use their phone or credit card. This ‘smile to pay’ function was introduced by Alipay as early as 2017 and by 2019 facial payment devices appeared in over 1000 convenience stores and more than 495 million Chinese people used facial recognition technology for payment in 2021 (Hu et al., 2023).

Chinese citizens also seem to be less concerned about the use of facial recognition technologies than citizens in other countries. According to an online survey, Chinese citizens only expressed low concerns about data collection by the state (Steinhardt et al., 2022: 1). Kostka et al. (2021) conducted online surveys in China, Germany, the United Kingdom and the United States and found that China had the highest citizen approval rate of facial recognition. The survey also found that citizens mainly associated facial recognition with convenience and improved security, rather than

with state surveillance and control. In our data, none of the customers expressed any surprise when they encountered the facial recognition kiosk.

The work of users in HCI

Our focus is on how the introduction of automation changes not so much but workers, but what customers do. Here we draw on research in HCI, in particular, studies that employ ethnomethodology and conversation analysis. The aim of such studies is to provide a contrast to other accounts of how such technology could or should work, by investigating instead what users actually do, which often includes focusing on unanticipated troubles that users encounter and solve.

In an influential study, Suchman (2007 [1987]) explored how users performed a series of tasks with a high-end photocopier. Suchman focussed on the troubles that users encountered and pointed out that a human observer could discern much more about the user's intentions than the machine could: 'the machine could only "perceive" that small subset of the users' actions that actually changed its state' (p. 11). On the basis of these observations, Suchman developed a powerful critique of the 'computational' or 'mathematical' models of communication. One of the innovations of Suchman's study was the fact that rather than simply asking how users interacted with the photocopier, she video-recorded these interactions.

More recent studies have also examined the detailed interactions between human users and smart devices based on detailed observations. Harper and Mentis (2013) conducted an ethnography exploring how family members played games with Microsoft Kinect. They investigated how users learned to interact with the machine and how they figured out what kinds of movements were, or were not, recognised by the machine. They showed that over time users adjusted their movements in order to produce behaviour that was 'understood' by the machine. In order to interact with the machine successfully, people could not behave 'naturally', but instead tended to produce movements that were excessive or exaggerated, resulting in the 'absurdity of movement that is required by the Kinect sensor' (p. 167).

In a video-based study, Pelikan and Broth (2016) investigated communicative problems occurring when users interacted with the humanoid *Nao* robot and how this interaction changes over time. They demonstrated that, through continuous interaction with the robot, users discovered what the robot was capable of, which prompted them to change their own speaking patterns, manipulating, for example their volume, the length of their turns (producing very short ones) and the words they used.

Porcheron et al. (2018) considered how people interact with the voice assistant 'Alexa' in home environments. Like other studies, this one highlights how users change how they talked to the machine over time. However, the authors situated these interactions in a social context and thus also showed how *other* people (those not talking to Alexa) had to change their behaviour by, for example being silent when others interacted with the device. This has also been studied by Lipp (2023) in a study of how users of care robots need to stand at a particular distance and use particular vocabulary in order to communicate with a care robot.

All of these studies have shown that in order to successfully interact with these technologies, users have to do various things to overcome the limitations of the technology. There is a lot of work necessary to make these automation technologies work. Rather than the machine understanding the users, it is often the users who have to learn to understand the machine and in turn change what they do in order to be recognised or understood.

Data collection

Fieldwork for this research took place in two self-service hotels of the same chain, which we call the 'L-hotel'. Customers make an online reservation in advance. When they arrive at the hotel, they unlock the hotel's door with their customised password, which has been sent to customers via text message. After customers enter the lobby (Figure 2), they self-register at one of two interactive kiosks in the lobby, by verifying their identity with facial recognition technology to access their room. Customers first need to place their ID card on the sensor of the machine to let the system access their personal data and match it with their reservation on the system. The kiosk will then start the facial recognition process. Once successful, the system will assign customers a room number and issue a password to enter the room.

We set up two GoPro cameras in the hotel lobby and collected data over a period of 15 months. In order to inform customers about our research, we posted information sheets and signs in Chinese and English next to the kiosks. These documents described the purpose of the project, the contact details of the researcher and information about the video recording, including mention of the fact that we might use non-anonymised images in publications. Following the recommendations of Heath et al. (2010: 18), the signs contained clear instructions for opting out of the study (e.g. say into the camera 'I don't want to participate in this project'), which happened nine times. Cases where minors were captured (three in total) were excluded from our collection. The 'opt-out' method



Figure 2. The lobby of the hotel with two self-registration kiosks.¹

for the video recording and the use of non-anonymised images were approved by the university research ethics committee.

We collected 674 cases of customers checking in. However, on 2 days, the facial recognition kiosk was not working, so all customers on those 2 days (51 in total) were approached by the service worker as soon as they entered the hotel lobby and never interacted with the kiosk. We were left with 623 customers approaching the facial recognition kiosk to check in. The recordings were analysed following standard practices of qualitative video analysis (Heath et al., 2010) informed by ethnomethodology (Garfinkel, 1967) and conversation analysis (Sacks, 1992). We started by watching all the cases in order to identify those that were particularly revealing about the registration process. These instances were then transcribed and analysed in detail. In this article, we represent cases in the form of a ‘graphic transcript’ (Laurier, 2014), using video stills that are either anonymised using different filters or rendered as line drawings.

Analysis

In the analysis, we follow the journey of customers as they enter the hotel, demonstrating how they are progressively disciplined by technologies throughout the interaction. We begin with customers discovering that no service personnel are present. We then describe how they pose for the self-registration kiosk and respond to the kiosk’s instructions. Finally, we discuss the emotional reactions that customers have.

Becoming aware of the machine

When customers arrive at the hotel, the first thing they must do is figure out how the process works. No doormen are available at the L-Hotel; the front door is locked and controlled by a code system. Thus, customers initially encounter a locked door and must find a way to enter the hotel on their own. Figure 3 shows one example.

In Figure 3, a customer is walking out of the elevator and attempting to enter the hotel. He finds that he cannot open the door directly (Figure 3[b]), which prompts him to look to the side where he notices a keypad (Figure 3[c]). He then turns to the keypad and sees a QR code above it. QR codes are extremely widespread in China as a means of accessing the official accounts of businesses, getting information and assistance, ordering food in a restaurant, as well as paying in a variety of settings, including street markets (He et al., 2023). Indeed, QR codes have been called ‘tech that changed a notion’ (Turrin, 2021, Chapter 5). Consequently, the customer knows that he needs to scan the QR code (Figure 3[e]). He then reads the information on his phone that details how to access the hotel (Figure 3[f]). He inputs his access code on the keypad (Figure 3[g]). The system broadcasts ‘Welcome to the L-Hotel’ and the door is unlocked, which allows the customer to enter.

This case illustrates the concept of the ‘working customer’. First, when entering an environment based around self-service, customers realise in the moment that no personnel are present to serve them (Dujarier, 2016: 561), even though they were informed of this when booking the room in advance. Second, customers must then undertake activities that were previously performed by service personnel (Glucksmann, 2016; Goodwin, 1988). In this case, the customer works to find ways to open the door himself.

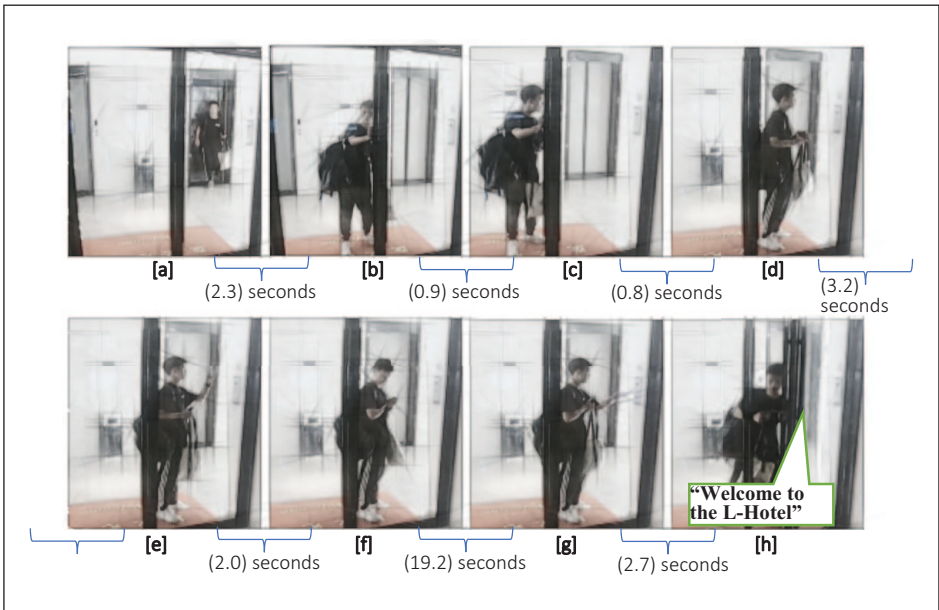


Figure 3. A customer enters the hotel (CU082).

Having now realised the self-service nature of the hotel, most customers immediately proceed to the self-registration kiosk that is situated directly ahead of the hotel entrance. Interestingly, they do so even in cases when they notice a service staff member sitting next to the door. In Figure 4, a customer sees a service worker (Figure 4[c]) but continues towards the facial recognition kiosk (Figure 4[d] and [e]).



Figure 4. A customer directly approaching the self-registration kiosk (CU036).

In rare cases when customers approach the service worker, they will be redirected towards the self-registration kiosk and asked to ‘do it yourself’ at the kiosk (Figure 5).

When customers try to enter the hotel, they quickly realise that it is not a ‘normal’ hotel; instead, they must work *with* their body to accomplish tasks (typing on a keypad, opening a door, etc.) themselves. However, when customers approach the

self-registration kiosk, they do something different: They start to work *on* their body in order to be recognised by the machine.

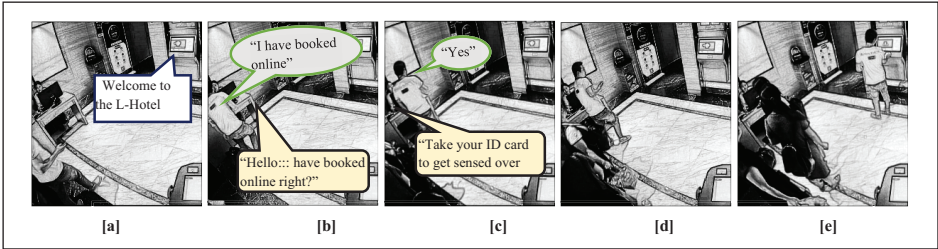


Figure 5. A customer approaching the service staff (CU022).

Presenting yourself to the machine

The self-registration kiosk verifies customers’ identity through facial recognition technology. When facial recognition starts, the machine will activate the camera, display a transparent ‘viewing frame’ on its interface (Figure 6) and announce the ongoing facial recognition.

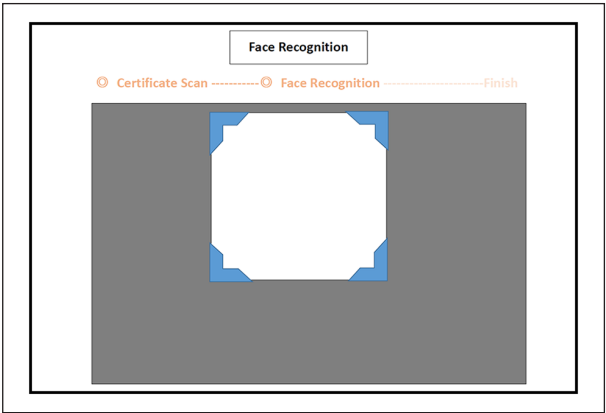


Figure 6. The viewing frame of facial recognition.

The first thing that the customers must do is ‘fit’ their face into the viewing frame. Figure 7 shows two examples. In the first, a female customer stands too close to the machine when the facial recognition begins and therefore moves a few steps backwards in order to position her entire head in the frame. In the second, a male customer moves closer to the machine during the process.

In the same manner as taking photos in a photo booth or taking selfies on their phones, the target (their face) needs to be placed within a particular area to be captured by the camera. Customers work to present their faces with the right angle and proximity.

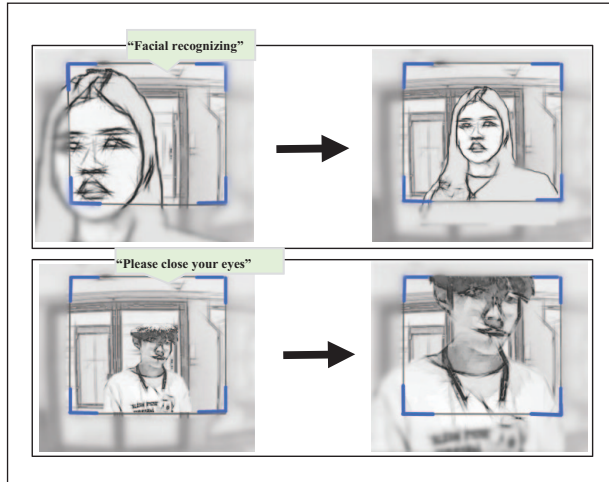


Figure 7. Customers adjusting standing position in front of the camera (CU028).²

Adjusting their standing position is not the only way for customers to make sure that their entire face is visible to the machine. They also ensure that there are no obstacles that could impede the recognition. Figure 8 shows how customers take off their glasses, hats, and masks for the facial recognition to succeed.

These actions reveal customers' common-sense knowledge of facial recognition technology: They are aware of the need to completely expose their face to the camera in order for their face to function as an 'index of their identity' (Gates, 2011: 8). Again, people are familiar with this process from other situations, such as passport control, where customs officers might ask passengers to remove masks or glasses to get a better look at their faces. In our setting, customers demonstrated their knowledge of facial recognition technologies by removing possible obstacles from their faces before and during the process.



Figure 8. Customers taking off glasses (CU118), hat (CU202) and mask (CU206).³

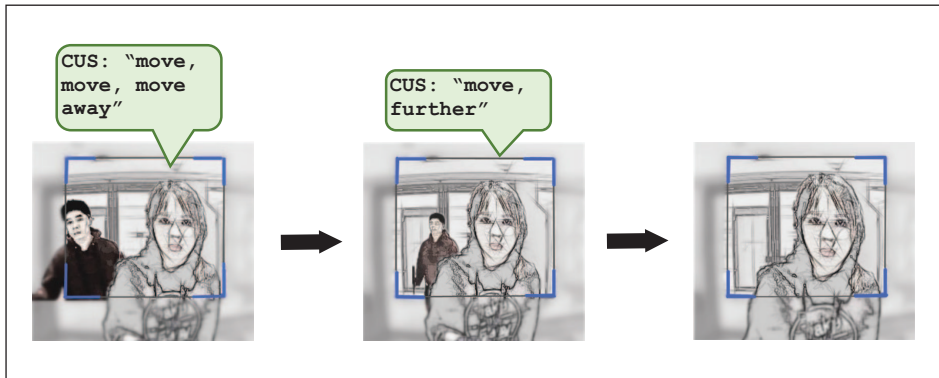


Figure 9. A female customer asks the male bystander to move away (CU123).

Customers using the facial recognition kiosk are aware that anything in the viewing frame can potentially pose a problem for successful recognition, including *other* customers. Customers often came in pairs. While one customer registered, the other could be captured by the camera as well. Figure 9 shows an example of a female customer asking her boyfriend to move away from the machine while she is performing the facial recognition process so that he disappears from the viewing frame.

Occasionally, bystanders (e.g. other customers or service staff) would offer explicit verbal advice on what the registering customer should do to complete the facial recognition process. We captured cases where the accompanying customer reminded the registering customer to remove accessories (e.g. ‘ah glasses’ [CU086]) or to face towards the camera (e.g. ‘look at the camera’ [CU249]). We also identified cases in which the service staff would suggest that customers take coverings off their faces and adjust their stances (e.g. ‘take off the glasses and step back a bit’ [CU124] or ‘stand one meter from the machine’ [CU147]).

When customers start the facial recognition process, they are no longer just working *with* their bodies; they are working *on* their bodies, standing in the ‘correct’ position and preparing the body appropriately. When one takes a photo of oneself, one makes an effort to look ‘good’; there is a negative version of the same reasoning, wherein customers try to avoid any problems that they can anticipate.

In these situations, customers are simultaneously a subject acting and an object to which something is done. This resonates with Bourdieu’s (1990) observation that a portrait ‘accomplishes the objectification of the self-image’ (p. 82) and Barthes’ (1982) comment on his photo-taking experience, ‘I am . . . a subject who feels he is becoming an object’ (p. 14). In other words, the customer becomes objectified into a *thing* during the human-machine communication. The self-registration kiosk prompts customers not only to work but also to manipulate their bodies—to discipline them—in ways that make them machine-recognisable.

Performing for the machine

Facial recognition involves more than simply taking a photo. The customers cannot stand still, as they would do for a photo, but must enact various bodily movements. They respond to a set of six instructions issued by the system:

1. 请闭眼, 'please close (your) eyes'
2. 睁眼, 'open (your) eyes'
3. 请摇头, 'please shake (your) head'
4. 摆正, 'stay still'
5. 请张嘴, 'please open (your) mouth'
6. 合嘴, 'close (your) mouth'

Figure 10 gives one example.



Figure 10. Customer responding to the machine's instructions (CUI38).

The technical explanation for these instructions is the process of 'liveness detection', which is a protection mechanism against 'spoofing', wherein a scammer would hold up a photograph rather than position a real person in front of the camera (Galbally et al., 2014; Pan et al., 2007). If the detection does not succeed immediately, the facial recognition kiosk goes through three 'cycles' of these six instructions, before it announces an error message ('Facial recognition failed') and prompts the customers to start another attempt. The liveness detection can succeed at any point. The earliest success in our data

occurred after three instructions (CU061), while the majority of customers succeeded during the first three cycles. Figure 11 gives an overview.

Attempts	Cycle	Number of customers who succeeded	Percentage	Cumulative percentage
Attempt 1 (A1)	Cycle 1 (A1C1)	102	17%	17%
	Cycle 2 (A1C2)	228	37%	54%
	Cycle 3 (A1C3)	194	32%	86%
Attempt 2 (A2)	Cycle 1 (A2C1)	63	10%	96%
	Cycle 2 (A2C2)	12	2%	98%
	Cycle 3 (A2C3)	2	0.3%	98%
Attempt 3 (A3)	Cycle 1 (A3C1)	4	0.7%	99%
	Cycle 2 (A3C2)	2	0.3%	99%
	Cycle 3 (A3C3)	1	0.2%	99%
Attempt 4 (A4)	Cycle 1 (A4C1)	2	0.3%	100%
	Cycle 2 (A4C2)	N/A	N/A	100%
	Cycle 3 (A4C3)	N/A	N/A	100%

Figure 11. Overview of the timing of facial recognition success (n = 610).

Out of 610 cases,⁴ 17% of customers achieved recognition during the first cycle of 6 instructions and 37% succeeded during the second cycle. The majority of customers (86%) were recognised during the first attempt, that is before the machine announced an error message.

The liveness detection means that customers do not just have to present their face to the camera they must also *perform* various bodily movements for the camera. This ‘performance’ is regulated and managed by the machine. While taking a selfie is something that people do for themselves, the movements made by the customers were done for the machine. It is the machine telling customers to close their eyes, shake their head and open their mouth. At a supermarket self-service checkout, customers follow instructions and accomplish various tasks *with* their body; here, they respond with appropriate movements *of* their body.

It is the third instruction that is perhaps the strangest. We often close our eyes (Figure 10, first row) for various reasons. We also may shake our head in public (Figure 10, second row) to, for example show disagreement. However, we rarely hold our mouths open (Figure 10, third row). Take, for example the second image in the third row of Figure 10. Out of context, it would be quite difficult to imagine a situation where that photo was taken—not during eating, but perhaps at the dentist. Figure 12 presents two more examples.



Figure 12. Customers' 'open mouth' movements (CUI18, CUI79).

When we present videos of customers performing these facial recognition tasks at conferences, audience members often smile and even laugh out loud at this 'open mouth' movement. To us, this reaction is evidence that there is something strange, perhaps even demeaning, about having to open your mouth for the machine. One reason for this strangeness may be that mouths are typically seen as a prime example of what Goffman (1971) has termed a 'territory of self', which comprises 'areas of concern' (p. 38). The recognition system objectifies every part of the body into a detectable element and routinises the mouth movement into the performance.

The customer must not only perform for the machine but also often *upgrade* their performance if it does not succeed. They are explicitly told that there is a problem after the machine has gone through three cycles of instructions ('Facial recognition failed; Please do it again' .), but they can become aware of a problem with their 'performance' much earlier: When they hear an instruction for a second time (during the second cycle). A repeated instruction is typically taken as a sign that the initial response to the instruction was in some way inadequate. For example, Macbeth (2004) observes that in school classrooms: 'Routinely, question-repeats are also heard as marks of a failed answer' (p. 719).

Socialised by their mundane experiences, customers perceive the appearance of 'problems' when they come across recurrent instructions. Consequently, the customers frequently modify their responses to the machine's instructions in subsequent cycles and attempts. Figure 13 shows how the customer changes how strongly they respond to the head-shaking instruction.

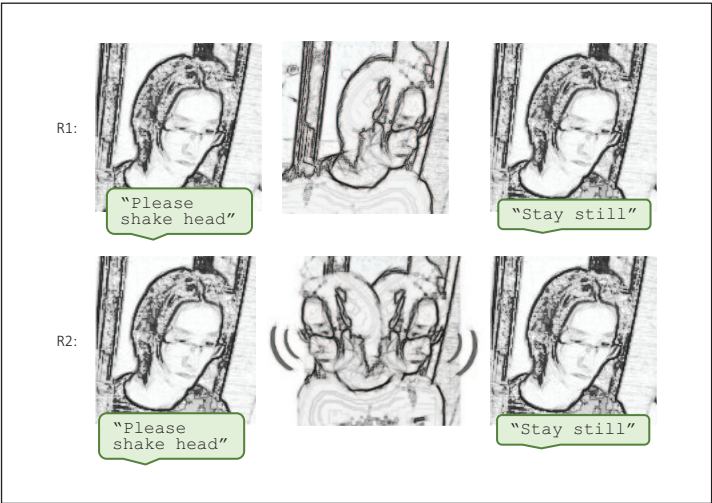


Figure 13. Customer shaking their head differently during the first and second cycles (CU035).

During the first cycle, the customer only turns his head to the left (i.e. only performs ‘half’ a shake). For the second cycle, he turns his head to the left and right twice (i.e. performs two ‘full’ shakes).

Figure 14 shows how another customer changes their response to the ‘please close your eyes’ instruction.



Figure 14. Customer closing their eyes differently during the first and second cycles (CU067).⁵

During the first cycle, the customer closes his eyes for only 0.3 seconds, before opening them (which occurs *before* the second instruction, ‘Open your eyes’, tells him to do so). In contrast, for the second cycle, the customer keeps his eyes closed much longer (for 1.2 seconds), waiting *until* he hears the next instruction.⁶

As Lipp (2023) details, one of the challenges of interacting with interactive machines is to deal with ‘ambiguity about what (non-)response or repetition on the part of the system means to the user’. In our case, customers engaged in a lot of remedial work in order to make the facial recognition process work (Greiffenhagen et al., 2023).

In this section, we have shown that facial recognition is more than just taking a photo. Customers have to ‘do’ something, and they have to do it as a performance for the machine, which involves disciplining themselves in various ways. Their eyes, heads and mouths are subject to the commands of the machine. In that way, customers are subjugating their bodies to the instructions of the machine, objectifying themselves in the process.

There is thus a stark asymmetry between the human and the machine, as it is the human who responds to the commands of the machine. This asymmetry is intensified in when problems occur. In human–human interaction, when there is a problem with the response to an instruction, it is typically the *instruction* that is calibrated (Goodwin and Cekaite, 2013). In other words, it is the instruction-giver who *modifies* their behaviour. However, the facial recognition kiosk will simply repeat the instruction again and again, so it is the response-giver who has to modify their behaviour, further enforcing the disciplining of the customer by the machine.

Expressing emotions towards the machine

Customers are aware that they are being disciplined by the machine, which can be illustrated through their emotional reactions during the facial recognition process. One of the most obvious emotions that customers displayed is frustration. In our data, customers occasionally become annoyed in response to repeated instructions and failures. Figure 15 provides two examples.



Figure 15. Customers swearing (CUI79, CUI19).

In the first case, a customer is annoyed when encountering the ‘please close your eyes’ order for the third time. In the second, a customer swears at the machine after it announces that the facial recognition process has failed. Cursing and impoliteness are common responses to communication with machinic interlocutors, especially when engaging with communicative failure.

However, customers do not only express frustration towards the machine. We also observed them smiling or laughing while following the instructions. For example, in Figure 10, the customer is smiling (or perhaps ‘smirking’) when following the instructions. Figure 16 features two more examples:



Figure 16. Customers smiling during facial recognition (CU256, CU083).

In the first case, the female customer grins when shaking her head. In the second, the male customer snickers at himself after the system instructs him to open and close his mouth. For us, these smiles, smirks and laughs are evidence that customers sometimes regard having to ‘perform’ for the machine as somewhat unnatural, awkward or embarrassing. The machine removes the customer’s control over their ‘presentation of self’ and disciplines their bodies.

Importantly, these reactions typically occurred when customers were not alone—when there was another human observing what they were doing ‘for’ the machine. If we zoom out from the second case in Figure 16, we see a female bystander watching the customer, which is displayed in Figure 17:



Figure 17. Bystander observing customer (CU083).

Studies have indicated that people usually feel relaxed when communicating with machines (Andrews, 2019). However, our cases illustrate that customers consider themselves to be losing face in front of others. This is similar to the situation described by Pentzold and Bischof (2023) in their ethnography supermarket self-checkouts where customers were often anxious, not so much of making mistakes, but of making mistakes *in front of others*, that is were afraid of ‘being watched while failing’. This was particularly the case if such onlookers were people they would know personally, which is exactly what we have in our setting. Indeed, bystanders often saw this as a funny performance. Figure 18 renders a scene wherein a bystander smirks at the registering customer, even mocking their actions:

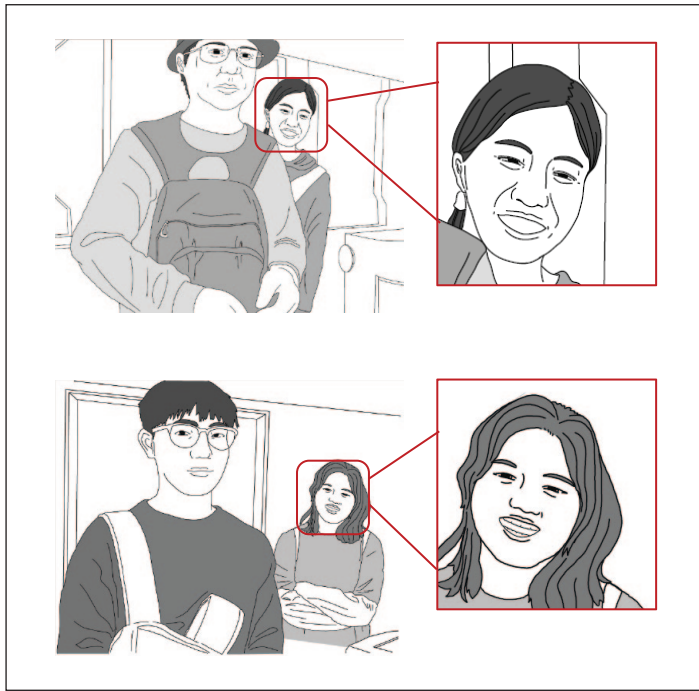


Figure 18. Bystanders smirking (CUI19, CUI38).

Compared with taking photos, customers using facial recognition have less control over what they should do. Though people objectify themselves in both practices, the interactivity varies between these processes. Objectification in photo-taking is relatively internal, positive and active, as people have the authority over how to present themselves in front of the camera. Objectification in facial recognition is rather external, passive and objective, as the machine dictates how they should perform. While people taking photos can do what they want and ‘be themselves’, people undergoing facial recognition need to react robotically to the machine’s commands. Such robot-like activities dehumanise the customer and transform the customer into another machine. The amusement of bystanders is made possible by the dehumanising of the registering customer, who goes from subject to object.

Discussion

In this article, we have shown that when customers are interacting with automated communication technologies, in this case, facial recognition kiosks, they are not just working; they are disciplined. Our notion of discipline indicates that customers not only conduct physical labour with their bodies when responding to the machine but also work on their bodies (indeed their selves) to become machine-recognisable.

Our argument resembles the one developed by Hochschild (2012 [1983]), who extended Marx's analysis of the labour of factory workers in the nineteenth century to service workers in the twentieth century. Hochschild argued that service workers are akin to factory workers, in that they have to perform physical work (such as preparing and handing out meals) and mental work (such as going through various steps in preparing for a landing). However, service workers must also do something 'more', which makes them different from factory workers. Hochschild called this additional work 'emotional labour', which 'requires one to induce or suppress feeling in order to sustain the outward countenance that produces the proper state of mind in others' (pp. 6–7). Throughout our analysis, we have argued that when interacting with the facial recognition kiosk, customers also do 'more'. This 'more' is not quite emotional labour—customers do not have to try to make the machine 'feel good', whatever that might mean—but it does involve a deeper engagement. In trying to be 'understood' or 'recognised' by the machine, customers alter how they speak, look and even act. They actively adjust where and how they stand, how their face is exposed to the camera (by removing various obstacles, including other customers) and, in light of repeated instructions, even how they close their eyes or shake their heads. They are no longer simply working *with* their bodies, but instead working *on* their bodies, which are being disciplined by the machine. All machines (such as ATMs or supermarket self-checkouts) control their users since users have to conform to the order to which the machine asks them to do something. However, in our setting, the technology actually 'talks' to users akin to a human telling them what to do. Furthermore, customers are not just asked to do certain things, but to perform in front of the machine in various ways.

This discipline may also lead to a new form of disenchantment or even dehumanisation. As discussed above, previous research has highlighted that the move from service to self-service comes with more than an increase in efficiency, predictability and standardisation—it also comes at the cost of removing personalised and customised experience for customers. Hochschild (2012 [1983]) likewise remarks that 'automation . . . leads to a decline in emotional labour, as machines replace the personal delivery of services, then this . . . trains people to be controlled in more impersonal ways' (p. 160). The introduction of automated technologies to the service industry thus is another example of Aneesh's (2006, 2009) 'algoratic control', which here does not subject workers to the discipline of the organisation, but subordinates customers to the commands of a machine. A key effect of this algoratic control is what Montagu and Matson (1983) term 'technological dehumanisation', 'the reduction of men to machines' (p. 9). Pentzold and Bischof (2023) describe how, when interacting with supermarket self-checkouts, shoppers are 'behaving more machine-like'. In our setting, customers not only perform movements according to the instructions and rhythm of the facial recognition kiosk but also render their own body

machine-like. They not only behave robotically, but become robot-like. What customers must do ‘for’ the machine is awkward, even embarrassing. Guzman and Lewis (2020) note that ‘AI technologies also raise questions regarding how people view themselves in light of their interactions with these devices’ (p. 78) and our research provides an example. Customers see that accessing a hotel no longer involves being treated like a ‘king’; instead, they must respond to the commands of a machine. Customers used to be greeted by the smiles of hotel staff, but they now (figuratively) smile for the machine.

Customers on occasion try to resist this. In supermarkets, customers may choose not to use the self-checkout (Andrews, 2019: 132; Pentzold and Bischof, 2023). In our setting, this was only possible when there was a technical problem (after all, the hotel was advertised as a ‘robotic hotel’). However, when things did not go smoothly, customers displayed a lot of defiance, especially in cases when checking in with someone else. This accounts for the swearing (Figure 15) of customers, as well as the amusement of bystanders (Figure 18).

Our case also raises interesting questions regarding the anthropomorphism of technology. As machines become more sophisticated, people often ‘anthropomorphise’ them (Duffy, 2003): they project social norms and interactional manners onto communicative machines, expecting them to ‘move’ as other human interlocutors. Our customers occasionally anthropomorphised the kiosk by calling it ‘motherfucker’ (CU119, Figure 15) or, in another case, ‘idiot’ (CU238). However, we also witness the opposite in this study. Since customers need to make themselves recognisable to the machine, they attempt to think and act like one themselves. Especially when the act is not effective immediately—when customers hear commands for a second or third time—customers need to figure out just what the problem could be *from the perspective of the machine*. They therefore make assumptions about how the facial recognition kiosk operates. This process is akin to the ‘technomorphism’ described by Vertesi (2012: 400–401), where she describes how the human team members of the Rover mission began to think more and more like the robot on Mars. In making themselves ‘machine-readable’ or ‘legible’, customers perform a technomorphic move, ‘seeing like a facial recognition kiosk’. Rather than endowing the facial recognition kiosk with human-like qualities, customers adhere to the logic of the machine and eventually act like a machine. There thus exists a danger in speaking of ‘human-machine communication’ as ‘communication’ might suggest that it takes place between two equal ‘partners’, each modifying their expectations and requirements for the other. At least in some situations, the interaction with automated communication technologies is not equal at all, and it is the human actor who subordinates their actions to the machine.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Part of this research was supported by a grant on “Work, Interaction, and Technology” from the Department of Applied Social Sciences at The Hong Kong Polytechnic University.

ORCID iD

Christian Greiffenhagen  <https://orcid.org/0000-0002-7544-9246>

Notes

1. Though it is possible for customers to register at the hotel completely on their own, for legal and technical reasons, there are still service workers present.
2. These are 'simulated' images based on our video recordings (which are taken from the side), designed to capture the 'viewing frame' from the perspective of kiosk.
3. These images are drawn from original video images.
4. This excludes 13 cases out of the total of 623 cases, in which customers approached the facial recognition kiosk, but were not able to complete the process. In three cases, the kiosk was turned off; in one case, the kiosk could not read the customer's ID card; in eight cases, the kiosk was able to read the card, but was not able to find the reservation information and in one case, the registration process started but froze before the facial recognition process was supposed to start. In all of these cases, upon encountering the problem, customers looked for a service worker who manually checked them in. Since the process either did not start or stopped working very early on, we did not observe customers expressing any particularly strong emotional reactions in these cases (in contrast to cases where the actual facial recognition process continued for a relatively long time—see next section).
5. Images are taken every 0.2 seconds.
6. This change occurred frequently in our data: After the first cycle, customers have learned that the instructions come in 'pairs'. Consequently, during the second and subsequent cycles, many customers keep their eyes closed *until* they hear the second instruction to open their eyes.

References

- Altenried M (2022) *The Digital Factory: The Human Labor of Automation*. Chicago, IL: University of Chicago Press.
- Andrejevic M (2019) *Automated Media*. New York: Routledge.
- Andrews CK (2019) *The Overworked Consumer*. Lanham, MD: Rowman & Littlefield.
- Aneesh A (2006) *Virtual Migration: The Programming of Globalization*. Durham, NC: Duke University Press.
- Aneesh A (2009) Global labor: algocratic modes of organization. *Sociological Theory* 27(4): 347–370.
- Barthes R (1982) *Camera Lucida: Reflections on Photography*. New York: Hill & Wang.
- Bourdieu P (1990) *Photography: A Middle-Brow Art*. Cambridge: Polity Press.
- Brush AB, Lee B, Mahajan R, et al. (2011) Home automation in the wild. In: *CHI '11: Proceedings of the SIGCHI conference on human factors in computing systems*, Vancouver, BC, Canada, 7–12 May, pp. 2115–2124. New York: ACM.
- Bucher T (2022) Facing AI: conceptualizing 'fAIce communication' as the modus operandi of facial recognition systems. *Media, Culture & Society* 44(4): 638–654.
- Bueno C (2020) The face revisited: using Deleuze and Guattari to explore the politics of algorithmic face recognition. *Theory, Culture & Society* 37(1): 73–91.
- Crawford K (2019) Regulate facial-recognition technology. *Nature* 572: 565–566.
- del Rio JS, Moctezuma D, Conde C, et al. (2016) Automated border control e-gates and facial recognition systems. *Computers & Security* 62: 49–72.
- Duffy BR (2003) Anthropomorphism and the social robot. *Robotics and Autonomous Systems* 42(3–4): 177–190.
- Dujarier MA (2016) The three sociological types of consumer work. *Journal of Consumer Culture* 16(2): 555–571.
- Foucault M (1978) *The History of Sexuality, Volume 1: An Introduction*. New York: Pantheon.

- Galbally J, Marcel S and Fierrez J (2014) Biometric antispoofing methods: a survey in face recognition. *IEEE Access* 2: 1530–1552.
- Garfinkel H (1967) *Studies in Ethnomethodology*. Hoboken, NJ: Prentice Hall.
- Gates KA (2011) *Our Biometric Future*. New York: New York University Press.
- Glucksman M (2016) Completing and complementing: the work of consumers in the division of labour. *Sociology* 50(5): 878–895.
- Goffman E (1971) Territories of self. In: Goffman E (ed.) *Relations in Public*. New York: Basic Books, pp. 28–61.
- Goodwin C (1988) ‘I can do it myself’: training the service consumer to contribute to service productivity. *Journal of Services Marketing* 2(4): 71–78.
- Goodwin MH and Cekaite A (2013) Calibration in directive/response sequences in family interaction. *Journal of Pragmatics* 46(1): 122–138.
- Goriunova O (2019) The digital subject: people as data as persons. *Theory, Culture & Society* 36(6): 125–145.
- Gray ML and Suri S (2019) *Ghost Work*. Boston, MA: Houghton Mifflin Harcourt.
- Greiffenhagen, C., Xu, X., & Reeves, S. (2023). The Work to Make Facial Recognition Work. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW1), 1–30.
- Guzman AL and Lewis SC (2020) Artificial intelligence and communication: a human–machine communication research agenda. *New Media & Society* 22(1): 70–86.
- Harper R and Mentis H (2013) The mocking gaze: the social organization of Kinect use. In: *CSCW ’13: Proceedings of the 2013 conference on computer supported cooperative work*, San Antonio, TX, 23–27 February, pp. 167–180. New York: ACM.
- Heath C, Hindmarsh J and Luff P (2010) *Video in Qualitative Research*. Thousand Oaks, CA: Sage.
- He C, He L, Lu Z, et al. (2023) ‘I have to use my son’s QR code to run the business’: unpacking senior street vendors’ challenges in mobile money collection in China. *Proceedings of the ACM on Human-Computer Interaction* 7: 1–28.
- Hepp A (2020) Artificial companions, social bots and work bots: communicative robots as research objects of media and communication studies. *Media, Culture & Society* 42(7–8): 1410–1426.
- He S, Zhang Y, Zhang J, et al. (2019) Why and why not use my face? A case study of face recognition solutions in the workplace. In: Marcus A and Wang W (eds) *Design, User Experience, and Usability: Practice and Case Studies*. Cham: Springer, pp. 485–504.
- Hochschild AR (2012 [1983]) *The Managed Heart: Commercialization of Human Feeling*. Berkeley, CA: University of California Press.
- Hu B, Liu YL and Yan W (2023) Should I scan my face? The influence of perceived value and trust on Chinese users’ intention to use facial recognition payment. *Telematics and Informatics* 78: 101951.
- Katzenbach C and Ulbricht L (2019) Algorithmic governance. *Internet Policy Review* 8(4): 1–18.
- Kostka G, Steinacker L and Meckel M (2021) Between security and convenience: facial recognition technology in the eyes of citizens in China, Germany, the United Kingdom, and the United States. *Public Understanding of Science* 30(6): 671–690.
- Laurier E (2014) The graphic transcript: poaching comic book grammar for inscribing the visual, spatial and temporal aspects of action. *Geography Compass* 8(4): 235–248.
- Lipp B (2023) Caring for robots: how care comes to matter in human-machine interfacing. *Social Studies of Science* 53: 660–685.
- Lukanova G and Ilieva G (2019) Robots, artificial intelligence, and service automation in hotels. In: Ivanov S and Webster C (eds) *Robots, Artificial Intelligence, and Service Automation in Travel, Tourism and Hospitality*. Bingley: Emerald, pp. 157–183.

- Macbeth D (2004) The relevance of repair for classroom correction. *Language in Society* 33(5): 703–736.
- Manokha I (2018) Surveillance, panopticism, and self-discipline in the digital age. *Surveillance & Society* 16(2): 219–237.
- Mateescu A and Elish M (2019) AI in context: the labor of integrating new technologies. *Data & Society Research Institute*. Available at: <https://datasociety.net/library/ai-in-context/>
- Montagu A and Matson FW (1983) *The Dehumanization of Man*. New York: McGraw-Hill.
- Munn L (2022) *Automation Is a Myth*. Stanford, CA: Stanford University Press.
- Palm M (2016) *Technologies of Consumer Labour: A History of Self-Service*. New York: Routledge.
- Pan G, Sun L, Wu Z, et al. (2007) Eyeblink-based anti-spoofing in face recognition from a generic webcam. In: *2007 IEEE 11th international conference on computer vision*, Rio de Janeiro, Brazil, 14–21 October, pp. 1–8. New York: IEEE.
- Pelikan HR and Broth M (2016) Why that nao? How humans adapt to a conventional humanoid robot in taking turns-at-talk. In: *CHI '16: Proceedings of the 2016 CHI conference on human factors in computing systems*, San Jose, CA, 7–12 May, pp. 4921–4932. New York: ACM.
- Pentzold C and Bischof A (2023) Achieving agency within imperfect automation: working customers and self-service technologies. *Convergence*. Epub ahead of print 17 May. DOI: 10.1177/13548565231174588.
- Porcheron M, Fischer JE, Reeves S, et al. (2018) Voice interfaces in everyday life. In: *CHI '18: Proceedings of the 2018 CHI conference on human factors in computing systems*, Montreal, QC, Canada, 21–26 April, pp. 1–12. New York: ACM.
- Pylypa J (1998) Power and bodily practice: applying the work of Foucault to an anthropology of the body. *Arizona Anthropologist* 13: 21–36.
- Rieder K and Voß GG (2010) The working customer—an emerging new type of consumer. *Psychology of Everyday Activity* 3(2): 2–10.
- Ritzer G (2013 [1993]) *The McDonaldization of Society* (20th Anniversary edition). Thousand Oaks, CA: Pine Forge Press.
- Robertson C (2009) A documentary regime of verification: the emergence of the US passport and the archival problematization of identity. *Cultural Studies* 23(3): 329–354.
- Sacks H (1992) *Lectures on Conversation*. Hoboken, NJ: Blackwell.
- Smith JE (2020) *Smart Machines and Service Work: Automation in an Age of Stagnation*. London: Reaktion Books.
- Steinhardt HC, Holzschuh L and MacDonald AW (2022) Dreading big brother or dreading big profit? Privacy concerns toward the state and companies in China. *First Monday* 27(12): fm.v27i12.12679.
- Suchman LA (2007 [1987]) *Human-Machine Reconfigurations: Plans and Situated Actions*. 2nd ed. Cambridge: Cambridge University Press.
- Turrin R (2021) *Cashless: China's Digital Currency Revolution*. Gold River, CA: Authority Publishing.
- Vertesi J (2012) Seeing like a rover: visualization, embodiment, and interaction on the mars exploration rover mission. *Social Studies of Science* 42(3): 393–414.
- Ziewitz M (2016) Governing algorithms: myth, mess, and methods. *Science, Technology, & Human Values* 41(1): 3–16.

Author biographies

Xinzhi XU is an Assistant Professor at the School of Journalism and Communication, Jinan University. Her research interests include Human-Robot Interaction, Human-Computer Communication, International Communication, computer-supported collaborative work, Ethnomethodology, and Ethnography.

Christian Greiffenhagen is Associate Professor at the Department of Applied Social Sciences, The Hong Kong Polytechnic University of Hong Kong, where he leads the Video Analysis, Science and Technology (VAST) Research Group, which develops video-based methodologies to study the impact of scientific and technological developments on people's lives. His recent publications include "Judging importance before checking correctness: Quick opinions in mathematical peer review" (*Science, Technology & Human Values*), "The Visibility of Digital Money: A Video Study of Mobile Payments Using WeChat Pay" (*Sociology*; with Li & Llewellyn), and "Orchestrated openings in video calls: Getting young left-behind children to greet their migrant parents" (*Journal of Pragmatics*, with Gan & Licoppe).