EPIC2019

Ethnographic Praxis in Industry Conference Proceedings

A.I. Among Us

Agency in a World of Cameras and Recognition Systems

KEN ANDERSON, *Intel Corporation*
MARIA BEZAITIS, *Intel Corporation*
CARL DISALVO, *Georgia Tech*
SUSAN FAULKNER, *Intel Corporation*

This paper reports on the use and perceptions of deployed A.I. and recognition social-material assemblages in China and the USA. A kaleidoscope of "boutique" instantiations is presented to show how meanings are emerging around A.I. and recognition. A model is presented to highlight that not all recognitions are the same. We conclude by noting A.I. and recognition systems challenge current practices for the EPIC community and the field of anthropology.

> Unknown, Caucasian, male, grey hair, 80 kgs, 1.8m, 55-60 years at entrance 2.
> Unknown, Caucasian, male, grey hair, 80 kgs, 1.9 m, 55-60 years in hallway 1.
> Unknown, Caucasian, male, grey hair, 78 kgs, 1.9 m, 55-60 years located in café 2.
> Unknown, Caucasian, male, grey hair, 80kgs, 1.8 m, 55-60 years located in hallway 3.
> Unknown, Caucasian, male, grey hair, 80 kgs, 1.8m, 55-60 years located in café 2.

Thousands of "observations" are logged, one about every second, during a single day on campus, ostensibly forming some sort of narrative of the researcher's day. What kind of narrative is it? That's the question. What the researcher understood at this stage was simply that this narrative was made possible by a set of networks of cameras connected together; a range of facial recognition systems dispersed across the school campus. Somewhere, or perhaps at multiple points distributed across the network, judgment and decisions were being made, that scripted the actions of others and thereby gave shape, unbeknownst to him, to the actions he might or might not take.

Strangers on campus are noted by the recognition software as "unknowns." This means that they are not students, staff, faculty, parents, administration, regular service people or even those identified as "concerns." By the end of a day visit, one of the authors had been spotted in the #2 café at least 3 times, usually in the company of another "unknown" and accompanied by someone who was known. This made the author a kind of "known unknown", which was an acceptable identity to the system, warranting no further action than to continue to register his presence. In this way, these school recognition systems demonstrated some small ability to deal with uncertainty. Looking from the camera's point of view, the author, and another researcher had become "familiar strangers" (Stanley Milgram,1972). Milgram used the concept to help explain the rise of modern cities. In this paper we are flipping it to help think about a new hybrid digital-social landscape being ushered in by A.I. and facial recognition.

BACKGROUND

Everyday life is a more mixed world experience than ever: digital/analog, machine/human, bits/atoms. Donna Haraway (1984) called out the limitations of such

binaries decades ago, and today such binaries are even more inadequate as our lives are even more hybrid, comprised of more-than-human multiplicities. Advances in artificial intelligence, cloud computing, wireless networking and data collection have ushered us into a new social-material era, one equally exciting and anxiety provoking. But relationships don't come easy and humans and technologies are surely in a protracted period of courting one another. If the industrial age ushered in one set of expectations and accountabilities, artificial intelligence seems to change the character of this courtship—suddenly our relations are much more promiscuous. In part, these distributed and varied encounters are expressive of a shift from products to networks, and concomitantly, a shift from discrete and singular artifacts of value to value as an outcome of connectedness and multiplicity. The shift is one where digital technologies that were previously limited to particular kinds of discreet, controlled, one-to-one interactions are now engaged in constant interaction with many, sometimes multitudes of humans. However, this adjustment period is the beginning, not the end. Self-driving cars, "personalized" agents on our smartphones and household systems, and autonomous robots are just some of the images conjured when A.I. is mentioned. While these examples seem to suggest A.I. is represented by a sleek, singular futuristic technological artifact, several scholars have highlighted how contemporary instantiations of A.I. rely on a complex, distributed, interdependent network of computers, software, data warehouses and infrastructure (Dourish 2016).

This paper offers a critical and ethnographically-informed exploration into key questions surrounding the constitution of A.I. and recognition systems as they permeate the complex practices and relationships that comprise contemporary everyday life. Our focus is on recognition, A.I. and the real time video analytics of recognition that are deployed and used in everyday contexts today. We will empirically illustrate the ways that human and non-human agents participate in building everyday life worlds and cooperate in this shared meaning-making process. We want to focus on the many agents involved, and shift the focus from singularity of device, product, service, and brand to the heterogeneity of intersecting databases, programs, products, services, people and networks.

We are conceptualizing various collections of A.I. and recognition as polyvocal *assemblages* (Tsing 2015, Deleuze and Guattari 2003, Ong and Collier 2004). The concept of the assemblage is salient because these systems are not in fact singularly engineered. They are diverse, more-than-human assortments that are gathered together, sometimes by design, other times *ad hoc*. Even though we might experience them through discreet interactions, as coherent services, their composition is multifaceted, often entangled. Our hope is to develop a critical appreciation for how diverse materialities, cultures, agencies, and experiences blend together in these emerging assemblages.

This use of assemblages has been employed to shift the framework of research to place greater emphasis on the dynamic, changing, and opaque characteristics of these A.I. recognition assemblages, as well as to bring in non-human participants. The approach enables agency of objects and the possibility of heterogeneity of assemblages. The researchers here are positioned to observe how elements are understood to cohere in existing or developing assemblages. Unlike Tsing's mushrooms (2015) or Bennett's (2009) green chilies, we did not have a material object to focus upon, rather this is the ground work to understanding how thoroughly entwined systems can mutate and develop over time [and space] and frame what is possible, desirable and expected of recognition systems. As Deleuze and Guattari (1987) note in their original writings on assemblage's, they are

"anticipatory" and concerned with continuing trajectories and future possibilities of what these assemblages might become, which seem particularly apt as we research A.I. and recognition technologies. The alternative of conceiving A.I. and recognition uses as discrete products or systems would imply a closed-ended and functionalist understanding that hides the series of interconnected and interdependent sets of technologies, institutions, agendas and people. What emerges here are partial directions and pressing questions related to the topic of the conference - agency: as artificial intelligence becomes an agent, what are the opportunities and challenges for shaping relationships to continue to enable agency? And what kinds of agency are possible in a world where technical things can know and do?

APPROACH

Since 2017 we have conducted four field research projects in China and two studies in the USA.[1] In 2017 we elected to study these A.I.-recognition technologies because they offered attractive solutions to address many contemporary needs for identification and verification. These technologies brought together the promise of other biometric systems that tie identity to individual, distinctive features of the body, and the more familiar functionality of video surveillance systems. This latter aspect has also made them controversial, which motivated our research to get a deeper understanding. In the USA, there has been growing social and political concern around the use of facial recognition systems. Samplings from the press in recent months include stories in the BBC (White 2019), Wired (Newman 2019), New York Times (Teicher 2019), Washington Post (Harwell 2019), CNN (Metz 2019) and The Guardian (2019), to name a few. In contrast, China's facial recognition systems, found in urban centers like Shanghai, Beijing and Hangzhou, were becoming ubiquitous even in 2017. In China, these recognition technologies continue to grow in sectors like civic behavior, retail, enterprise, transportation and education. Business Times (2019) reports that Alipay facial recognition payment is already deployed in 100 cities and will pay $582 million to expand further. Tencent, is adding facial recognition payments to the WeChat platform of 600M users. In a society that has had overt and everyday surveillance in human and institutional form for over 70 years, the emergence and deployment of recognition through cameras has been less controversial than in the USA.

We also chose to study these systems because recognition technologies, for all of their social and political controversy, allowed us to continue to talk about humans. Unlike some other A.I. systems, recognition technologies rely upon human embodiment, action, and often interaction. This is significantly different from, for example, machine learning systems that use social media as proxies for human activity. We hypothesized early that camera systems were harbingers of new interaction models with humans, and that recognition technologies, in particular, were examples of cameras literally reaching out to people, albeit awkwardly and often inaccurately. For even when deployed as a surveillance use case, the experience of being seen at a distance in a public space equipped with CCTV was a kind of interaction that implicated a more complex web of human users with specific interests and motivations. These new interaction models are suggestive of notions of embodied interaction (Dourish 2001) but also, due to the seamlessness of these recognition systems, these new interactions also seem to elude some of the situations of collaborative meaning-making we are accustomed to. As these systems become so commonplace that they disappear, and our interactions with them become just another everyday action ("smile to

pay"), how do we—humans—participate with these dynamic, but elusive assemblages to make the worlds we want to inhabit?

In 2017 facial recognition systems were emerging in the mainstream landscape at a global scale just as companies like Intel were shifting business interests to the cloud and networks, and in the communications arena to 5G. The technologies emerging to transform the network, mobilized further by 5G's emphasis on machine to machine compute, indirectly signaled that the interaction model of human and device, a hallmark of the PC ecosystem, was no longer the asset to exploit. Today's technology industry conversations about "edge" and the challenge faced not just by silicon companies, but by cloud service providers, telecommunications companies, telecom equipment manufacturers, original equipment manufacturers, and even content providers on "last mile access" and how to bring compute closer to where data is produced, simply do not focus on what people do with technology. In this business context, increasingly distant from end-users, facial recognition provided us with a way to continue to talk about humans at a moment where so many only wanted to talk about machines.

Finally, we were skeptical not about the fact of facial recognition becoming ubiquitous in China, but about the contrast cultivated by the USA press relative to deployments at home. The research concerns in the USA on facial recognition have centered on three points: 1) recognition systems were biased in their development (Burrell 2016; Crawford and Shultz 2013; Eubanks 2017; Noble 2016; O'Neil 2016; and Pasquale 2016); 2) the systems created new risks to privacy (Dwork and Mulligan 2016; Introna 2009); and 3) there were ethical concerns about use (Horvitz and Mulligan 2015; Stark 2019). While Eubanks (2017) has equated their development to the rise of "eugenics", Stark (2019) equates the potential dangers of recognition to "plutonium." But these concerns have not necessarily resulted in fewer systems adopted. Indeed, Gartner (Blackman 2019) projects recognition to be the fastest growing Internet of Things (IOT) space in the near future. Further, we have seen deployments expand in the USA since 2017 in public city infrastructure as well as airports, private school campuses, industrial facilities, summer camps and childcare settings. Further, the US government says facial recognition will be deployed at the top twenty US airports by 2021 for "100 percent of all international passengers," including American citizens, according to an executive order issued by President Trump (2017). By examining deployed uses of recognition, we hoped to provide empirical evidence to fill the gap between building, speculation and future deployments.

In what follows we share a kaleidoscope of vignettes from the field to supply the raw material for a discussion about value and its complexities for A.I. and recognition. The use of kaleidoscope is intentional in that it is not the scientific instruments of telescope or microscope that we employ here, but images of instantiations of new technology with people; images left open for further interpretations. As Gibson (1999) notes, "The future is already here – it's just not evenly distributed." While there has been plenty of speculation on the cataclysmic possibilities of A.I., there has been a dearth of studies on tangible, instantiations; so, something that is more "what it is" than "what might it be." We will share snapshots of a future world of A.I. and recognition that is already here. We focus on what could be called "intimate" or "boutique" uses of recognition; so, not massive surveillance systems, but closed institutions or community uses. The snapshots don't tell a complete story--there isn't one to tell--nor do they provide a perfect compass for navigating the emerging new spaces unfolding before us. Instead, they are glimpses into the kinds of

questions a compass can address, and the kinds of terrain it should help us navigate. From these vignettes, we raise questions about future research and practice for the EPIC community.

STORIES FROM THE FIELD

Everyday & Uneventful Facial Recognition

Popular visions of A.I. are seductive, but real-world facial recognition is amazingly boring in China. A few of the A.I. systems we experienced delivered identification for seamless access to residences, offices and schools; seamless access to subways and trains; seamless identification for hotel check-in, and seamless access inside banks and at the ATM; clerk-less convenience stores; preferential treatment in retail stores; identification for government services and criminal investigations. This list of the applications is only meant to underscore that A.I. and recognition is commonplace in China, and still growing in both government and commercial sectors, to the extent those are differentiated. From the start, what is important to emphasize is how banal the use of these systems is. Perhaps there is complexity and prowess behind the scenes, but everyday interactions with these systems and services is…well…every day.

Recognition is so ordinary and uneventful that it often goes unnoticed, both to users and to researchers who are supposed to be in the field keenly observing. As a result, there were *many* times in the field when we had to ask people to repeat their use of a facial recognition system, so we could observe the process. We asked one of our early participants in the study if we could take her picture as she walked through the facial recognition system at her residence. She walked through, and we had to ask her to do it again. We explained she did it too fast for us; that we could not see the system in action. Could she do it again? Ooops, we missed it the second time, and then we missed it again the third. Finally, we just asked her to walk *very* slowly, much slower than usual, and we got it. Of course, by that time a mother and her kid, an older woman, and the security guard were all looking at us like we were idiots. The guard, in particular, seemed delighted by it all. Another time, there was the look of a young man when we asked to go with him to take money out of the facial recognition ATM. You could almost see him thinking, "Oh yeah, foreigners think facial recognition is interesting? Is this a scam to take my money?" We also had to ask him to log in three times to catch the process.

Such interactions with facial recognition are very different from, indeed opposite to what we are used to with technology. Generally, with any kind of technology, whether a personal computer, phone, Alexa, Nest thermometer, car, or even Siri—we *prepare* to interact, and we remain aware of the interface, even with those that work almost seamlessly. Facial recognition interactions in China are stunning because they are so normative and normalized, often blending seamlessly into the environment. For example, three women walking back into work after lunch only briefly look in the direction of the facial recognition machines as they continue to walk and talk straight back into the building. Nothing to see here. No break in the conversation. Hardly a pause in their steps. They give a look that is less than a nod one might give a security guard that you knew very well. It is substantially less of an action than pulling out a badge, and pausing to badge in. Life simply unfolds, not only as if the technology was never there, but also as if those social regimes and routines of

observation that define so much of what we call society and culture had ceased to exist. But of course, that haven't ceased to exist, they've just been differently delegated.

Facial recognition is not just a part of high-end office buildings or residential complexes or trendy businesses; it is becoming commonplace everywhere in China. We watched as customers at a KFC quickly ordered on a screen then smiled briefly to pay. Yes, giving up money and smiling about it! In practical terms, of course, the smile is a second form of authentication for the facial recognition system to verify that you are alive (first the system verifies you are you; smiling is a secondary measure to avoid spoofing). The "smile and pay" is also common at some grocery stores. "Sometimes you can't help but feel a little happy about smiling [even if it as a machine]" a woman checking out at a grocery store commented. Of course, she isn't really smiling at a screen. She is smiling at an Alipay system (from ANT Financial) that is part of the Sesame Credit loyalty program for Alibaba. People are aware of the Alibaba loyalty program, and some of the perks of participation. Dual systems, like the ticket/person verification system at the Beijing main train station are also popular, as lines move quickly with people being recognized, authenticated and verified by a machine, rather than waiting in the lines to get tickets and then waiting for a security person to check in before boarding. These are just normal, everyday, "nothing to see here" parts of urban life.

Beyond the mundaneness of recognition systems, people were able to articulate some advantages, and while they would raise occasional issues about use, their concerns did not necessarily impinge on the value of using a facial recognition system. People mentioned that it is more secure, is hassle free because all you have to do is smile to get access, and oh yeah, it is fast. On the surface, these seem to be values of efficiency — where ease of use and enhanced productivity determine the worth of the system. While that may be partially the case, we also believe users found meaning and significance in the fact that the use of these systems removed and obviated the unnecessary social complications often inherent in transactions. In other words, one of the (human-centered) values of these systems is the desire to avoid awkward interactions with other humans in a socio-cultural context that has weighed heavily on how those interactions should take place. While social interactions are important in China, they come at a cost. People may push more stuff at you to buy or try to make connections by attempting to leverage a transaction into a relationship. There are additional cultural factors at play here, such as those of class. Though we presume people want to interact, and that sociability is desired, that presumption may be flawed, or at least not always true or uniform. By their very personalization, recognition technologies support the capacity to elide select social encounters.

Participants in the study were expecting to see more places and more uses for facial recognition in their urban environment. Unlike the USA, there was no moral panic, in fact, people were excited and proud about what they perceived to be a highly novel technology.[2] There is a solid cultural belief in China's middle class that technology is both a marker and a catalyst for economic growth and national success on the global stage. The recognition systems are interpreted as markers of the development of society, at the same time they are making urban China an easier place to live, and in some respects more like the West. In a curious way, A.I. facial recognition technologies highlight the individual, a hallmark of Western culture and traditions. As one of the participants said, "If everything is connected then you can just bring your face!"

Someone Is Watching You: Interpretive Flexibility

High School X: Hall Security[3]

High School X, in a tier two city in China, has switched their campus security camera system over to one that uses facial recognition. The facial recognition system enables students to come and go freely on campus and is connected to the classroom attendance (check-in at the door) system. The security camera system can be accessed from any authorized desktop, e.g., security office monitors, IT office PCs, principal's PC, etc. The school used to have a bank of twelve TV monitors rotating through the twenty cameras on campus. The campus now has over forty cameras on campus for security. Two features of the system were demonstrated for us. One feature of the system was that it does anomaly detection of spaces and, when possible, identifies the person in the space (minimally captures them). Anomaly detection in this case means someone is in a space at the wrong time, e.g., in the hallway during class time. The other feature enabled a human supervisor to search by image or name in order to have all the appearances of that person for the day aggregated on screen. Taken together, these capacities enabled the detection of more than just attendance. As the following example shows, they enabled the detection of patterns of behavior, and as a consequence, revealed relations that might otherwise go undetected.

> [Interview 1PM Classroom]
> June (HS X Student): I've had cameras in my schools all of my life. They are watching us to protect us, but it is a little creepy. I mean, they know so much about us that they could know when you go to the bathroom or if you were dating, and who that is, really anything . . .
>
> [Interview 3PM IT office]
> Main IT guy (HS X): I think you talked to June earlier. Did she mention she was dating? Dating between students is not permitted at this school. We've known[with the facial recognition system], she has been dating for over a month. We haven't done or said anything about it. She and her boyfriend are both getting very good grades. As long as they are getting good grades and don't disrupt the community (school body), we won't interfere.

How did IT and the administration know June was dating? We don't know. Those details weren't forthcoming. We do know that the analysis of her daily patterns involved verification with a teacher, the anomaly detection, and person identification (like a game of Clue) on the school grounds. The interpretive agency in the assemblage didn't reside solely with the software but with the interaction between security, IT, teachers and the hall monitoring software.

Cindy Toddler Monitoring

Cindy is raising her two toddlers in Shanghai with the help of two nannies, her in-laws, a cook, and seven in-home surveillance cameras. Cameras in almost every room are used to monitor activities and behaviors, to understand when a routine is broken, to look for lost items or to trace the root cause of a dispute.

Cindy operates a centralized system where her children are the assets and she is the processing hub. All the analytics run through Cindy who uses the cameras to collect data she uses to monitor and investigate activities in order to shape the behaviors of other actors responsible for her children's care. In one incident described during our fieldwork,

> Cindy goes home to find her son and nanny are napping earlier than the established schedule. Cindy reviews the camera footage to understand what transpired and sees her mother-in-law fighting with the nanny who proceeds to retreat to the bedroom with her son. Cindy understands the context for the earlier nap time and reprimands her mother-in-law via WeChat text. When the nap is over, Cindy instructs the nanny in person about mother-in-law best practices.

In Cindy's system, the data inputs may be distributed, but analytics and decision-making are centralized. Her system's performance requires a particular set of members (nannies, parents, in-laws) to align to a particular set of values and practices (regarding food, hygiene, sleep, play) that demonstrate her version of good parenting. Cindy taps her system of cameras to access data and make sense of the actions and events that do and do not follow protocol. This constantly updated contextual insight allows Cindy to intervene and correct the behavior of the other human actors as needed to maintain optimal performance.

St. Nicholas School Safety (USA)

A similar situation unfolds at St. Nicholas of Myra, a private Catholic Pre-K to 8th grade school in a gentrifying urban neighborhood. The principal at St. Nicholas of Myra has recently deployed a facial recognition system. The recognition system is made up of humans, multiple cameras and computer technology. The cameras at St. Nicholas of Myra are used to monitor who comes in and out of the school and "to know the community better." Unlike either of the HS systems in China, the system at St. Nicholas of Myra only identifies adults, not students or anyone under eighteen. The principal and receptionist see a face and name on the facial recognition system monitors for almost every adult including the milk delivery person and the food staff. This allows the principal and the school receptionist to make sure the right people have access to the school. The system allows the principal and receptionist to identify and greet everyone by name, which they feel fosters a feeling of community. The principal sees his role as making sure the kids are "safe, happy, healthy and holy," and feels the facial recognition program helps him to achieve those goals.

Ways of Watching

Of course, the staff at HS X, Cindy, and the Catholic school principal actively manage how people act and exert power in their respective systems; a fact that is not dependent on the presence of cameras. They do so in the name of particular kinds of human value, but there are key differences in how that value is produced because cameras are present. In Cindy's case, value lies in her ability to care for her children the way that she wants through resources she has enlisted (nannies, in-laws, etc.). For Cindy, value is achieved by restricting the capacity of her nannies and in-laws to act independently of her parenting plans and goals, plus introducing the capacity of the camera to document what has taken place. In doing so, Cindy uses the camera as a means of witnessing, producing evidence that she employs, to

ends that are of her own choosing. Indeed, the camera data gives Cindy another partial view on what took place—not the nanny's or her in-law's. Cindy's understanding, enabled by the camera, allows her to shape the human links between herself and her nannies and between the nannies and her in-laws ("best mother-in-law practices"). This human work doesn't disappear; rather the presence of a camera enables it and gives Cindy more direct control over it. Conflicts may be deviations from the plan, but they also give Cindy the opportunity to work on stitching together human relationships that are central to the system.

In the St. Nicholas of Myra case, monitoring access and movement in the school increases social connectedness and an overall sense of community, but does not prevent all bad things from happening. If an unknown person or a person marked by the system (entered manually by the principal) as a "concern" tries to enter the school, the door will not open unless the receptionist or principal unlocks it. For instance, a parent suffering from substance abuse who is not currently allowed to see his kids, will be blocked by the system from entering the school. Here the opportunities for mistakes or misuse are rife, but trust is placed in the principal to make these decisions—extending his capacities to act, but still allowing him to retain authority over the system.

In China's HS X, school administrators guard against disruption to the learning environment from both inside and out. The disruption can be at the individual or the community level. Anyone not granted access is blocked, just as in the St. Nicholas of Myra system. But this system is more proactive in monitoring internal activities. Kids skipping classes, rough housing, regular visitors going places they aren't authorized to be, are all behaviors that can lead to a decision to act. Previously, if one of the same people had noticed an irregularity, they would also act. This resembled the system at St Nicholas of Myra, where the principal or receptionist using the camera monitoring system can spot kids hanging out under a main staircase in the school – a place they shouldn't be during school hours. One key difference is that the camera system brings the situation to the immediate attention of security, or others if they are on the system, so action can happen sooner. The other key difference is in the ability to pull together a series of incidents over time; to create a narrative of what took place. Sam, a student at HS X, was known by the system of technology, security, IT and administration, to skip class occasionally, after checking in on the camera system. He would go out to a remote (unmonitored) part of the garden area on campus, smoke, read books, and work on his homework until the class session ended. They knew he did this because they could see him out of class and entering the garden on video. Security people learned about the smoking. None of that was acceptable behavior generally, but because Sam was one of the top students in his class and did nothing that would hurt or infringe upon his classmates, this was permitted. They school officials were willing to assume that Sam just had days when he needed to get away. The principal at St Nicholas of Myra made similar kinds of decisions when he spotted kids hanging out under the stairs, for instance. He wondered, is this just a kid trying to disappear in the midst of a bad day or are kids engaged in improper or destructive behavior? In both cases, humans continue to own the judgment about the importance of the behavior. Based on a calculation of value, they are willing to interpret and to read between the proverbial lines to explain the student's behavior beyond what policy permits. Staff or teachers can then speak to the students about their behaviors, and so create new paths for human to human interaction. The human work doesn't disappear, but is enabled, managed and focused by the cameras.

Agency Denying Systems

Steamed fish today. No chips.

Chinese High School Z had a nutritional system that was powered in part by facial recognition. It was really not "a system," but five independent projects built upon each other: cafeteria ordering system, cafeteria and cafe payment system, cafeteria delivery system and two different vending machine systems. Besides incorporating different applications, there were at least three different recognition software pieces integral to the system, so even the core underlying programs were not shared. When we visited, all the food a student could acquire on campus was nutritionally noted to generate a recommendation for eating. Based on what the student had eaten, the nutrition was evaluated, scored and recommendations sent to the HS administration, and the student, and the parents. The student could then determine what, if anything, they might change in what they selected to eat. However, the system was doing not always work to enable student-led decisions.

Initially, the school ran the system so that the student would have a meal at the cafeteria that was predetermined, based on a student's optimal nutritional in-take. If the student's optimal nutritional in-take exceeded the guidelines on one day, the system would compensate and adjust the guideline to be nutritionally appropriate on the following day. A student could order whatever she wanted as long as it fit the guidelines. In practice this meant that students whose nutritional intake was deliberately constrained might get served steamed fish in the cafeteria instead of the barbecued pork. These same students might have their access to one of the vending machines blocked. Students who mapped to the need for guidelines had virtually no agency to select their own food since the system would make value judgments and constrain decisions on their behalf.

This food selection and decision-making system for students lasted less than a month. Parents and students both complained fiercely ("after all we (parents) we're paying for the food so our son should be able to choose what he wants"). Parents suggested to school administrators that the school should have a nutritional system similar to Sesame Credit where it would offer rewards, not punishments so students could earn points for special foods, or credits for the vending machines. The HS Z didn't have a way to economically implement this type of system. Today, the system is designed to enable conversations. It provides students with a view onto how they are doing, from a nutritional standpoint, for the day and for the week, and on how their behavior, indeed performance, matches to the suggested standards from the government. Parents can encourage their kids to eat correctly. They can have conversations with their kids about the administration's idea of how they should eat. Although, in the course of our research, we did not encounter any stories of parents who reported having those conversations with their kids. Finally, the students can use the report as a guide to reflect on food choices.

With respect to the cases that we observed, China's recognition systems do not appear to be bad things. The nutrition systems, at least in one case, was redesigned to help to bring awareness to some choices, actions and behaviors; awareness that could be used to adjust behavior towards desired goals. These examples show us that recognitions systems go wrong when they act alone to deny options to humans, who have their own creativity, ingenuity and agency to solve problems. The nutrition system as it operates today has been reduced from an active agent that determines what food is consumed to an off-site coach. The lack of

malleability or flexibility for the students in the initial system created a brittle partnership which did not get traction with students or parents. Students were not learning new skills. Parents were frustrated with unseemly distinctions. Both sets of stakeholders were constrained by a system, rather than encouraged to work with it. In China, this sort of system failed.

Personalize It!

Students, teachers, administrators, parents, and even IT people in the schools all talked about the hope that A.I. technology in the schools would increase personalized learning. Squirrel A.I. Learning, a private, A.I.-powered tutoring service in China, had become fairly well known as an after-school program using A.I. to generate personalized drill and practice sessions to improve students' scores on national tests. The public schools didn't have a computer per child to replicate that kind of personalized A.I. program. However, they did have cameras in classrooms. One camera set-up was tasked with taking attendance during class and it worked well. In addition to knowing who was in class, the parent-faculty-IT-admin community thought the camera and A.I. could create a better learning environment to know how the students were feeling, and in particular that it would recognize when they were "confused" "bored" or "frustrated" in class. [ENDNOTE 3] The IT-admins contacted a company to build an experimental system for them, though this didn't work out satisfactorily. The company said it could deliver an attention system that could tell whether a student was paying attention in class or not. Given that a typical class size is around fifty, this was perceived by the school as a way to ensure each student was engaged with the work (and so going to do their best). It would give the teacher insight into which students he or she was able to engage, or not able to reach. Because the key goals of the system were to 1) help students to learn more and 2) improve teacher performance, the system was assumed to cater to all classroom stakeholders. Further, for students and administrations, this would be a means to assure "no teacher bias" in the process of helping the students, or as American's might say, no favoritism in how attention is distributed to "teacher's pets."

The company provided the hardware and software. The system had two A.I. components, a facial recognition component and an affect detection component. The facial recognition was tied to the student ID data base. They guaranteed a 97% accuracy on affect detection, on the specific dimension of attention. The system had one camera mounted at the front of rooms that did an S scan every minute. The system would recognize each face and deliver an "attention" value (yes/no). Nested up at the top of a wall, it was virtually invisible, near to the camera that took attendance.

The teacher had a live report of the class activity (bottom of screen) and an overall report on the class session on his/her computer screen. The teacher was expected to be able to respond in-class to adapt their lesson in order to better engage the students. Students and their parents were sent a report with a percentage assigned to the dimension of "attention" in the class session. The students were supposed to try to improve their overall attention towards the teachers in class in the next session. The administration also had access to the reports on the class session for both students and teachers.

Parents started to complain within a couple of days about "privacy" violations of the system. At a different school there had been leaks of video footage of classroom activity by one of the school's camera systems. Some of the footage was humorous or embarrassing to

some students. Some parents were concerned that video moments when their child was "inattentive" would be caught and "escape" onto the Internet. The system had other problems that were working against it. Although no one disputed the facial recognition part, some felt uncertain that what the system "thought" and what their child was "actually" doing were at odds. For instance, some parents argued that, "My son concentrates with his head down on the desk. He is paying attention not sleeping," because they feared their child's behavior would be interpreted as inattentive. While verifying a student's identity (matched to photos) was perceived to be a straightforward process by parents and students, determining attention was perceived to be an inference. It was subjective. The affect detection technology may have had high accuracy in some dimensions, but it wasn't accurate in the way the community thought it should be. The school community discovered that it needed a human agent, such as the teacher, to interpret the data and then to take some immediate action, both for effective interpretation and action. The roles in the assemblage needed realignment. The school community learned an important point: that A.I. recognition assemblages are all probabilistic, never 100% accurate. They introduce a new kind of interaction with computer infrastructure that isn't about 0/1, right-wrong, correct-incorrect, etc. because by definition A.I. will always be wrong at some point, in some circumstance. The community's solution was to propose to increase the presence of the human agent in the assemblage to help negotiate value for the teachers and students.

All of these insights result in too much complexity to deal with. The affect detection experiment was quickly shut down.

> The affect experiment did not work . . . we learned a lot . . . we expected too much from the technology and not enough of ourselves. . . . we'll continue to experiment with new ways to help students & teachers in schools. . . . We're exploring a system that can detect actions like reading, writing, raising hands . . . That might come before the next affect use - HS Principal.

The community came together to shut down this system. The system did not have a life beyond what its constituents enabled it to have. Social forces prevailed. The teachers, administration, parents and students' still believed in A.I. recognition technology, and felt it would eventually lead to a better learning environment – a win-win for everyone. The path forward, however, was clearly going to be one of experimentation to enable more learning in the slow process of people forming new relationships with the technologies. "There may never be a perfect system, but we can do better," said one of the IT people involved in the set-up. The community, however, still had agency to put a stop to the recognition technologies, as well as, to be actively engaged to create what the next recognition technology should be and do.

Perfectly Imperfect: A.I. Is Human Too

Many of the particular systems we have discussed—eating, attention--have been part of larger systems, for instance as extended means to create better learning environments. One of the systems we explored in the USA was the use of facial recognition by a sheriff's department. What is striking about this context of use is the lack of agency the facial recognition software is granted, and conversely, the ways in which human agency is retained.

This might not be surprising were it not for the amount of agency such law enforcement facial recognition applications are believed to have based on repeated, reports about police departments use of facial recognition leading to bad results (Brewster 2019; Einhorn 2019; Garvie 2019; Stat 2019; and White 2019). Facial recognition applications were deemed so bad that San Francisco (Thadani 2019)and Oakland (Ravani 2019) have banned use by police departments and Portland, OR (Ellis 2019) is considering it.

For the Sheriff's Department of Rock County, facial recognition software is used in a very particular way by one particular department: as a partner in a larger more distributed crime solving team. The sheriff and detectives collect video of a crime. In the case highlighted in our research, they collected video of a theft that had occurred at a local store. Sometimes the video comes from neighborhood cameras, other times from other stores' security cameras, and still other times, from both. In this case, the footage was from an in-store camera. The guidelines for the sheriff's department are very clear in that the video does not come from any city or county public cameras, it only comes from private residential or commercial cameras. Often the video from these residential and in-store cameras isn't good enough quality to be used with the sheriff's department system.

Once the video is acquired, detectives work with the agency's Special Investigations Unit using facial recognition software to see if an image of the perpetrator's face from the store's surveillance footage is a match with an image from the internal database of convicted criminal mugshots from the county system. An algorithm makes a template of the face, measures the shapes of features and their relative distances from each other. A database consisting solely of convicted persons' photos from the county is then searched as the source of potential candidates — not photos from the Department of Motor Vehicles, not Facebook, not traffic cameras or the myriad streams of close-circuit TV video from around the city. What's more, facial "landmarks" are compared without reference to race, gender or ethnicity.

After the software generates a list of possible matches, an investigator assesses their resemblance to the suspect. Typically, there are 5 multiple hits. There is nothing visible to the investigators on the accuracy of the hits—it is simply a list of 5 previously convicted individuals who might be a match for the person in the video. The county realizes that the system is not perfectly accurate. Sometimes, the team decides none of the mugshots is a correct match. If one or more is selected, a review is conducted by detectives and supervisors, noting similarities and differences. If a person is selected from this list, that person becomes an investigative lead. The identification team will provide only a single lead to the case detective. If they affirm a match, the detective proceeds with further research, pursuing it like any other lead they would get, e.g., an anonymous caller, witnesses at the scene, 911 call etc. Notably, no one can be arrested on the basis of the computer match alone. For an arrest to happen, there must be traditional verifiable evidence of probable cause for an arrest. As such, the photo match does not count as legal "evidence." The facial recognition system is "just one input among many in our 100% human driven investigations" said one of the identification team members. His colleague added, "it provides a simple solution to an otherwise-tedious hunt through photos." And while the facial recognition doesn't count as evidence, the investigators see it as at least as reliable a lead as some eye witness accounts.

Other police departments in the USA have tried to give facial recognition systems more power in the police force, as is the case in Orlando, but they have been shut down (Stat

2019). Raji and Buolamwini (2019) examined all commercial facial recognition systems in the USA and highlighted the flaws and inadequacies of the systems in addition to fundamental injustices perpetrated by those inaccuracies. The assumption in these understandings of the facial recognition systems is that they need to have closer to perfect accuracy, operate independently of humans and have trustworthy value. This sheriff's office is an interesting case in that it assumes the system isn't perfect, just as the sheriff's deputies aren't perfect, and so sets in place a series of procedures to account for [non]human frailties. Technology–human interactions are frequently reduced to being thought of as issues around trust. Trust seems inaccurate to describe the role facial recognition technology is playing. The system has the accountability to discover the suspect, and because the system has many agents in it this accountability is necessarily shared. The 'black boxing" (Crawford and Schultz 2013) of the recognition system, or the investigator, or the detective, or the eye witness, etc. is not crucial as it is part of a distributed system of action.

FRAMEWORK FOR THINKING ABOUT RECOGNITION SYSTEMS

We have demonstrated a range of uses of A.I. and recognition assemblages. While still new and "cutting edge", it seems clear to us that these systems are rapidly becoming a commodity infrastructure that even small businesses will be able to build new applications upon. Across the research, we identified seven variables that give us a way to start to account for how these assemblages work and when and why they stop working:

Explicit permission. Does the agent give permission to be part of the system and know? Is it voluntary? Is the person aware of what is being recognized and why? Or is the hidden and unclear?

Recourse – is the path to correct any problems clear and reasonable. Recognition is probabilistic, which means at some point it will be wrong. Knowing this, having an actionable course of action when things are not right is important;

Consistent – is the system deployment consistent with the institution's stated business interests?

Personally Efficient – is the system deployment easy and does it achieve something of value for those being used as data. Of course, there can also be some broad community value (e.g., community health or safety). Or even more distant, the recognition is generating value for some other entities benefit;

Anonymized –are the data anonymized? is any personal identifiable information necessary to participate? Is it possible for the system to deliver personalized results if the information in the system is anonymized?

High Confidence – all recognition systems are probabilistic, though some are better than others and some instances are more difficult to determine. This measure looks to whether the use case will have high confidence or a high threshold in determining the result. At the extreme other end would be a system that requires human agents to make a determination.;

Self-contained –does the information stay within one domain or does it leak out to other domains, (e.g., residence access recognition isn't used in any other way and stays within the resident community's system)?

What follows is a brief introduction that applies some of these variables to show how the different assemblages using recognition software are distinct. We'll provide three examples to help draw out the differences between these variables, how they work and how they work together.

HS Access Facial Recognition

Our HS X used facial recognition to allow people (students, faculty, admin etc.) onto the grounds. The access set-up is very explicit and obvious. People give their permission to be part of it or if they opt out, they can use their ID cards to enter (albeit a slower process). If they are not recognized and blocked from entering, then they can see a security guard in a nearby booth and pass through with an ID. Knowing who is or who isn't on campus is considered part of the school's responsibility to students, staff and parents. By simply walking into school, it has eliminated long lines and wait times as people used to have to show their ID cards to guards and if their ID cards were lost or misplaced, it turned into an ordeal for people and the administration. There is no anonymization. The location and time of the person passing are noted for the daily records. There was high confidence that the recognition system would work since the data base was less than 1000 people. The data base and the results were contained to the school system only, which was an on-premise system. The mapping onto our vectors can be seen in Figure 1.

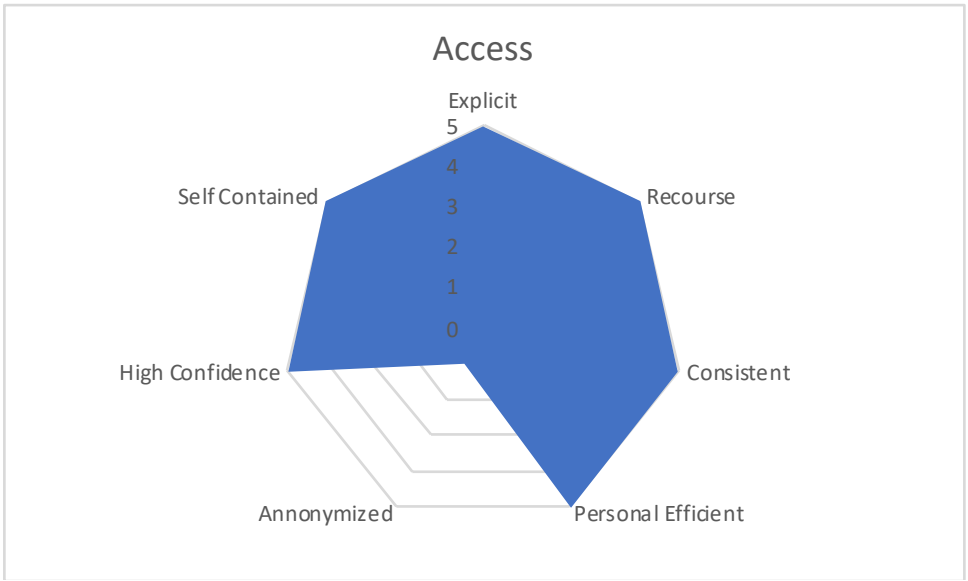

Figure 1. Access to School Facial Recognition Mapping

HS Affect Detection

Affect detection, though taking place in the same context, a school, has a very different profile than access to the campus grounds (Figure 2). While the explicit permission to be part of affect recognition might on the surface appear similar, it varies from the access example because the cameras are mounted up and away from the students. Because the cameras scan the entire room, one is never sure when they are being monitored. There is little recourse to the affect result – neither the student nor the teacher can know when affect moments happen, so they can't be contested or corrected. Because the classroom experience is about paying attention to the teacher, people felt it was an appropriate thing for the school to try to work to improve. While in theory there was value to the student and the teacher, neither was actionable value. The net result ended up being uncertain value for everyone. The recognition was directly tied to identified individuals who were given reports. The quality of the data set for what constituted attention/not attention , as well as, how behaviors were interpreted, was highly suspect. Video was accessible off campus by parents and the partner company.

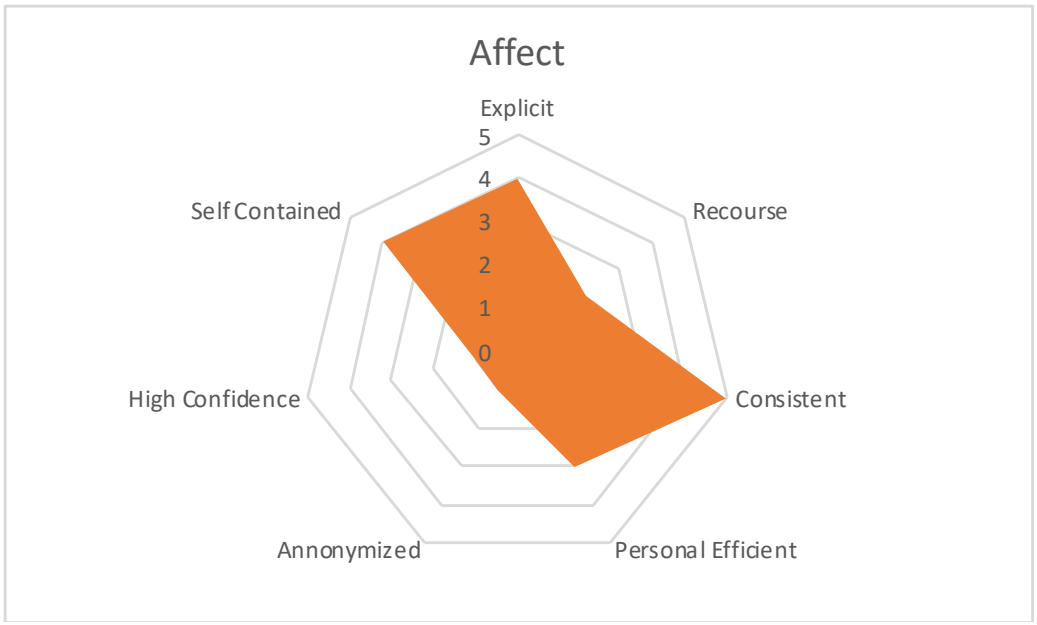

Figure 2. Affect Detection In Class Mapping

HS Hallway Recognition

Hallway recognition creates a slightly different profile than either of the above (Figure 3). While it too takes place in a school, it has a very different profile. While the explicit permission to be part of a hallway recognition might on the surface appear similar, it varies from the access example because the cameras are often mounted up and away visually from the students, almost hidden. There is little recourse relative to the hallway detection result –

moments are collected, but not necessarily immediately acted upon. Counts of activity can be made, without the video being retained. The IT person and/or security person have the dominant voice in interpretation. While administrators and teachers felt the system was consistent with the schools goals (safety, attendance, & learning), many students understood the safety and attendance aspect but felt the school should primarily be concerned with campus access and what happens in the classroom. The students did not see any personal value to the system. Overall the community value was insuring no misbehavior on campus creating a safer social and physical environment. There was no anonymization of the data – data was tied to an individual or individuals. It was recognized by all participants that both the recognition of the individual and of the activity were subject to a lot of interpretation by IT and security personnel. The hall recognition system was contained to the school environment with security access given only to particular people with particular roles in that school.

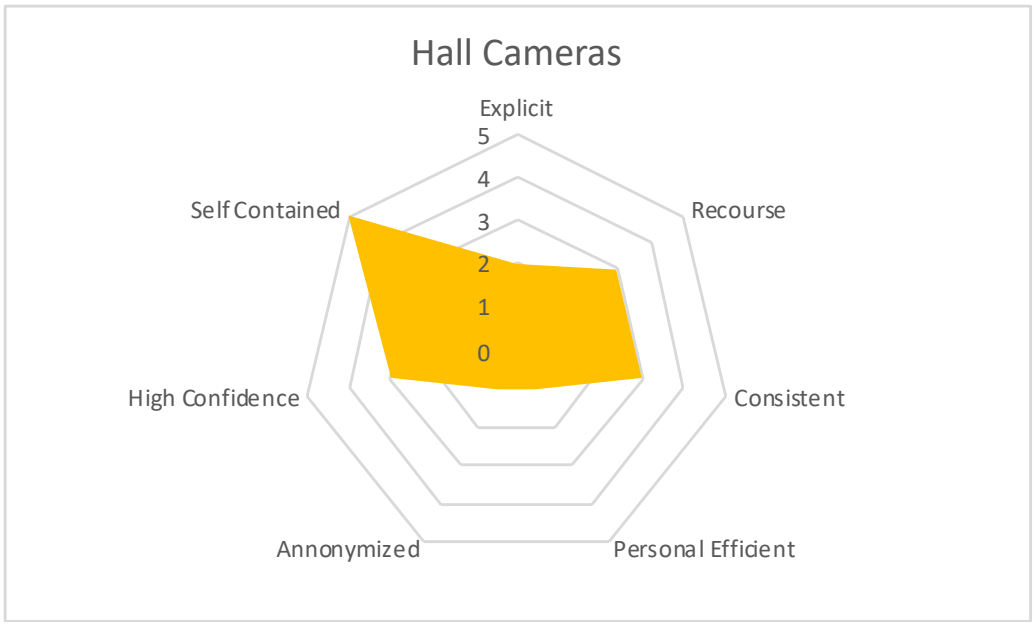

Figure 3. Hall Cameras in Recognition Mapping

While the diagrams provide a "systems approach" to think through recognition technology uses for those we have discussed and others that might emerge they are ultimately incomplete models. Specifically, these models do not address the important differences between A.I. (instructions, intentions, revealed preferences, ideal preferences, interests and values) on an individual or a collective basis. A challenge remains for researchers to identify fair principles for alignment on recognition technologies that receive reflective endorsement despite widespread variation in people's and communities moral beliefs.

DISCUSSION

All of the assemblage involving recognition software described here can be cast as providing for well-being, broadly construed (or at least that the intention by those who use them). The form and content of well-being differs from instantiation to instantiation, in some cases they seek to provide security, in others, health, in others a sense of comfort. In many cases, these forms are swirled together. They strive towards a holistic environment or milieu, characterized by values and desires that are projected into and through these systems. Surveillance is offered as the tool, the means to achieve that well-being. This is not, in fact, such an odd perspective. Regimes of observation, inspection, and supervision have long been part of how we, as individuals and societies, work towards well-being, whether through a disciplinary gaze or an ethics of self-care (Foucault 1995). What differentiates these regimes is the assemblage that enacts them and with which that we interact. Contemporary assemblages, such as the recognition systems we've discussed, display (if not possess) agencies of their own, capacities to act and exert power in dynamic ways that are new and unfamiliar. This requires that we do more than extend the existing theories of observation and control onto these assemblages. This requires that we work to articulate new theories that engage the agentic capacities of these assemblages.

These agentic capacities are apparent in the tailored character of these assemblages; the well-being generated is not generic. The aim of these assemblages is a well-being that is personalized in ways that people find meaningful. The subjectivities of the consumer are different from those of the citizen, which are different again from those of the student. These subjectivities are also always intersectional—the Chinese mother and the parochial school principal are complex inter-weavings of the social. Personalization then is more than a surface acknowledgment of the differences between one individual and another in order to deliver recommendations that cater to demographic differences. The rhetoric of personalization in an age of A.I. is about new sources of everyday benefit and fulfillment, enabled by new types of partnerships that bring new types of distance and intimacies into our relationships with other humans and with technologies; partnerships that help us to produce the worlds we want to inhabit. Of course, we can and should question this rhetoric, but the point remains, personalization in the age of A.I. is not the transactional customization of Web 2.0.

While the research represented here is limited, the socio-material change in the definition of "the field" brought about by recognition systems strongly suggests the need for new or modified approaches for doing innovation work. We see at least three aspects of our work that could be (re)considered: 1) assemblages, not individuals or user experience; 2) where we get our models for A.I. networked systems; and 3) the necessity of a humanities approach.

Assemblages, Not Individuals or Groups

As a community of practice, we should consider a shift in our lens from the individual experience to the collective, technical, institutional, and regulatory systems that surround peoples who exist in networks of assemblages. Studying "users" as we have traditionally conceived of them will be of limited help in understanding the transformations that A.I. and

recognition will enable or force in society. Our familiar ways of thinking and working are likely to limit themselves to the failures of a particular instantiation of a particular system in existing socio-technical contexts as we know them. But this will not be helpful for understanding the contexts that are emergent from A.I. assemblages.

It would be a failure to think about the principal or parents or students or teachers or security staff or IT personnel as being the only generative actors here. The technology, government, markets and institutions create affordances that enable particular kinds of agency, which in turn interact with those technologies. Ethnographic traditions like those that emerged following Geertz in anthropology or The Chicago School, like Howard Becker, in sociology, wanted to account for the larger frameworks that guided action and understanding (cultural in the first, social in the later). Following in those traditions, we see, for example, the user plus the direct user experience plus the use of one or more A.I. programs plus the policies of the Chinese government plus market forces (implicating companies like Hikvision, Intel, Alibaba, Baidu, etc.), as well as incentives around efficiency (what we think machines could do) – all as part of what we've referred to as the A.I. and recognition "assemblage." In this context for research, the individual user, or for that matter, even the notion of a group, should be re-case as an assemblage, which encompasses all of those who use or would be affected by the use of the system, imbricated with multiple cultures, practices, institutions and markets. We do this not by forcing us to see how this stuff affects individuals, but how this stuff *is* the assemblage.

In the end, the importance is not that the A.I. has its own agency, nor that users make A.I., but that A.I. is making new kinds of people, individuals and society (among other things).

Some might suggest that existing methodologies, like Actor Network Theory, offer this opportunity. While such methodologies are a potential starting point, what's really needed are methodologies that enable us to be more anticipatory of how value might be created, and less analytical of how valuation has already occurred. For instance, as we partner with these systems, we need to develop an appreciation for new modes and experiences of agency. Agency has never been reducible to the capacity for human action alone—as if people were ever able or willing to act independently of the worlds they make and inhabit. Capacities for action and exerting power are an outcome of an intermingling between people, other humans and a multitude of other things. Agency is a quality and effect of networks. Here, Actor Network Theory is a useful starting point. ANT posits that what we consider to be the social world is a constantly shifting series of relationships of humans and non-humans of varying scales that function together (Latour 2005). What is distinctive about this method is that it does not privilege humans within the network. Agency is not a quality of any individual actant but rather of the configuration of the network. As that configuration is dynamic, so too are the agencies within that network.

Another important aspect of agency within ANT, which distinguishes it from many other perspectives, is that agency does not require intentionality. So, for instance, in discussing the issues of restocking a bay with scallops, it is fair to describe the ways in which the scallops themselves are actants and refuse to participate in this process (Callon 1984). Such a flattening of subjectivities and ontologies is disturbing to some social theorists, but precisely the point of ANT: to de-center the human and consider an expanded perspective on how the world is made and then made to work. Proponents of ANT are quick to point out that ANT is less a theory of the social and more a method for tracing the associations and processes by which what we call the social comes into formation and actions. Given its

attention too, indeed its embrace of, heterogenous collectives of humans and non-humans, ANT has proven to be particularly useful for the description of contemporary conditions in which objects and systems regularly are taken to be acting in and on the world.

But ANT alone is not enough. In fact, ANT may not be the most useful starting point in a world populated by A.I. algorithms and socio-technical networks. ANT is an analytic tool that allows us to describe the world, after it has been made. It is less useful for understanding the world *as it is being made*, and perhaps totally unhelpful as a framework for making the world as we might want it to become. What is needed are practices and theories that enable us to better imagine how the world might be made—concepts of networks and agency that help us to explore the distance and intimacies that we have to deal with today; concepts of networks and agency that are imaginative, exploratory, and speculative but also grounded in fundamental humanistic principles based in the possibility of relationships.

Contexts as Models Of and For – Beyond the Literal

While we considered many different A.I./recognition systems as they were being deployed, we were reminded of a key direction for innovating new communication and information systems, that is by researching those that have been around for hundreds of years. This is a radical departure from traditional research for what has become classical UX and innovation work that looks first at the immediate and literal context of use as a site for product/service intervention, followed by work on ever more specific requirements for said product or service. If you are creating a product for baby food or travel mugs or working on how to make a better Xerox machine, this may have been adequate. But communication and information assemblages may or may not be modeled in the intended context and the variables that need to be contextually informed have more to do with data flows than actual sites of use. An alternative in the innovation process could be researching cultural contexts and systems that can illustrate the data flows and exemplify the goals of the system to be designed. In short, some research needs to take place outside of literal context in order to find its actual context.

So, if you want to create an A.I. recognition system that might get used in a stadium or an autonomous vehicle, looking at the actual context of use may not necessarily be the best place to ground the research. Instead, exploring a site that has characteristics of a robust and intelligent network might generate new ways of thinking. For example, researching the medina networks in Morocco may provide new ways of thinking about the kinds of resources that computational networks will have to make available. In these markets, we can see how tourist networks learn to interact with existing networks of vendors and local guides. These kind of research sites might provide a better model for a smart network or pulling together an assemblage, than looking at the actual classroom, where that same technology in question is meant to be deployed. Human systems are incredibly innovative and time-tested and are often ignored as "systems" and reduced to literal contexts, actual contexts of use. To paraphrase Geertz, we shouldn't be limited to creating models *of* some particular context of innovation but also models *for* innovative systems. Separating the models for design from contexts for implementation invites new perspectives and frameworks for innovating complex assemblages of solutions.

The shift from individuals to assemblages, the changing character of what we once referred to as context also suggests that, as a community, we need to broaden the theories

and methods we engage in, while also parting ways with techniques that no longer serve us. While there is an ongoing need for researching human cultural contexts of use, there is a limit to what we can understand by observing the use of these systems by people, in part because so much of the system itself is not encountered by humans in use. To better contribute to a vibrant imagination of how the world might be made, we need to complement our practices of observation with practices of interpretation. Thus, another implication is the need to draw theories and methods from the humanities to better understand these systems. What do the humanities offer? Certainly, more than empathy. What the humanities offer are ways to interpret the things that humans make—"readings" of many kinds, close readings, distant readings, reparative readings, deconstructive readings, and so on. These readings are also designs in the sense that they are acts that organize ways of life, ways of living in the world. They provide a critical lens into the systems that claim to produce meaning and even knowledge. Importantly, these acts of reading are fundamentally different than observing what humans do. We tend to think of the humanities as providing skills for the interpretation not just of poems, literature, paintings and such, but of video games, logistics systems, algorithms and new categories of texts that provide the means to be human in a more-than-human world. To develop a fuller appreciation for what these systems are, and might be, we need to continue to develop practices of ethnography in an expanded field, which recognizes the need for, and the limitations of, human-centered in a world comprised of artificial intelligence, and looks to bring practices of interpretation to the fore.

In addition, recognizing the limitation of how we study these systems and their contexts of use, we should also acknowledge limitations on how we communicate our research. The techniques and tools of representation we have used in the past seem worn and shredded as we take on these dynamic assemblages. Many of these techniques and tools were developed in the context of human factors, in the context of designing interfaces for systems in which there were material affordances or the ability to create facsimiles of material affordances. What is more, most of these techniques and tools place emphasis on the individual and their interactions with a system that is bounded. But as we've discussed, that is simply no longer the case. It is not enough to tell the story of a system from the perspective of a single person, or a single product, and it may not even be enough to tell the story of a system from a human perspective alone. Personas are likely inadequate to capture a recognition program. A use case fails at articulating the value, dynamics, and complexity of education in the classroom. How do we tell stories that are polyvocal, wherein some of those voices are not-human? How do we represent dynamic configurations of agency?

CONCLUSIONS

We have presented glimpses into a subset of processes in which social realities are becoming realized in and around recognition assemblages. These glimpses start to show how it is that verbs of doing become nouns of being (to watch, am watched). It is a start on a longer pathway of discovery on how our lived worlds are pragmatically produced, socially construed, and naturalized. In many ways, A.I., beyond ML, is still so abstract, diffuse, and unknown. In this paper, we have tried to shift the conversation from the potentially soteriological or cataclysmic possibilities of A.I., to what is firm, clear, steady, and tangible; moving beyond just something that is more "what might it be" than "what it is." Rather than

considering A.I. hypothetically in all of tomorrows futures, our interest has been to examine A.I. as it is instantiated, experienced in practice and culture today. Only by capturing moments now, are we able to understand how A.I. among us is creating new kinds of individuals, institutions and society.

In the end, there are many questions about what exactly are the problems in contemporary A.I. systems for social sciences and how to investigate them ethnographically. It is not as if the social sciences are just coming to A.I. —there are decades of work to build from on social-material systems. And yet, out contemporary A.I. systems seem to be distinct in the ways humans are instrumentalized for the sake of nonhumans. The human action is material for the nonhuman algorithm. The kinds of assemblages that A.I. is bringing together challenge us to consider what our practice is and how ethnography matters in it. Are projects studying the engineer working on algorithms in a cube or software teams in a lab going to be enough? Anthropology started as a study of "man" <sic> the animal, in an evolutionary and comparative framework. Today, we are shifting over to an understanding of people in a cybernetic framework; an understanding of people as machines with nerves. New instantiations of A.I. challenge us to consider what it means to be human, or nonhuman. It pushes in a direction complimentary to "multi-species" ethnography (Kohn 2013) or anthropology beyond the human (Besky and Blanchette 2019). These new A.I. instantiations also suggest new ways to frame and do our work. Considering possibilities of following data flows, like Mintz (1985) did with sugar, or considering assemblage subjectivities, instead of just individual ones. To understand the implications of these assemblages to the human, we have to better understand the nonhumans. The anthropological project around post-human This requires experimentation new ethnographic techniques (Seaver 2017).

With this massive and yet occasionally quiet shift slowly but surely taking place, we have the opportunity to reflect on our roles as corporate social scientists, humanities thinkers, ethnographers, design researchers. We have choices to make about the degree to which we will continue to work to improve the technologies, services and assemblages that continue to expand the role of A.I. in our daily lives, or if we will work to slow down the rate of adoption, in some cases, going so far as to argue against it. Neither these technologies nor our study of them is neutral. While we should remember that we've been here before—with the invention of electricity, automobiles and even television—we recognize that A.I. systems and assemblages are different, more invasive, and place into check values and principles that humans have claimed for themselves. It's another crossroads for our applied disciplines and our shared interest in ethnographic work. Perhaps instead of posing the options as binaries—as choices we each need to make to advance one option at the cost of the other—we can work to improve *and* to slow down and in doing so to recognize that these two paths more than likely coincide at every step.

NOTES

Acknowledgments – We'd like to thank all the people who gave us time out of their busy day to share their thoughts, stories and experiences with us. We'd also like to thank the institutions that opened doors for us and welcomed us into their communities. Finally, we'd like to thank Ellie Rennie, our EPIC curator, for her truly helpful comments and support.

1. We will draw upon research primarily from China with some comparative or contrastive sites in the USA. Pseudonyms are used throughout this paper. The research in China was conducted in 2018. We spent two weeks surveying recognition programs in public use in Beijing, Shanghai and Hangzhou. Primarily these were one on one around particular recognition programs, e.g., access to banking, access to work, smiling to pay, etc. While trying to understand how the systems were used (and others they used), we also explored the broader context of their lives. We returned 6 months later and spent 10 days to do deeper dives around recognition systems in educational institutions. We primarily focused on 3 high schools: 2 public and 1 private. The schools discussed in this paper are both public schools. One school was one of the poorer ones in the district, while the other was situated in a university community. All the recognition systems discussed were not yet commercial systems. At the schools, we interviewed a variety of stakeholders: teachers, administrators, staff, students and parents. Independent from the interviews at schools, we talked to representatives of some of the companies that provided the systems to the schools. The school administration asked that their schools names not be used in any report. Likewise, all the participants in the research have been anonymized. None of the systems created for the schools in China were products or services at the time we did our research – they were experiments. High School Z uses a team of parents, teachers, staff and administration to brain storm uses for new applications that they want to bring onto campus. The administrator and IT lead try to find (large or small) companies interested in creating the system for the school, creating public and private partnerships. The public schools in China, in general, when we were in doing the research, had no guidance for systems to build, buy or deploy – everything was an experiment. The research in the USA was primarily site visits. We visited the sheriff's department in May of 2018 and the St Nicholas school in March of 2019. The facial recognition software used by St Nicholas is a commercial product. The former was done as a part of the exploration of landscape of uses of facial recognition. The later was conducted as a point of comparison to what we had seen in China.

2. When we were in China, the stories about facial recognition systems being used on the Uyghurs had not become content of mainstream media in the USA or China. The stories of facial recognition that were circulating were about people being ticketed for minor offenses (e.g., jay walking), dispensing toilet paper, and criminals being identified and/or caught on the street (or at events), authenticating appropriate car service drivers and so on. The camera surveillance system was primarily explained in terms of safety and civic etiquette, reinforcing the way people were to behave, protecting against those who violate etiquette and laws. No one we talked to wanted to see less recognition systems in place, most had ideas of where they wanted to see more, e.g., "ticket dog poopers who aren't scoopers" "find my child" "reward appropriate behavior in Starbucks (throwing trash away)."

3. As mentioned, the recognition systems in schools should be considered experiments. The affect system was an experiment to create a better classroom experience for learning. For those in the USA, the in-school experience is a little different, particularly when looking at something like affect detection. The value of the student is judged more on how he/she/they perform on the national exams then on grades in school. Every class I saw, someone slept during class. The reason given was they had been studying non-class material for the national exam until late in the night and were tired. All students and parents talked about the use of materials from outside of the school work to help them with the national exam. The import of the exam vs the school plays out in the various systems in that the evaluation of the system about the student (attentive or not) does not really impact the student as much as such a system might in the USA. Of course, everyone wants to score well on everything, however, whereas a grade in a course might greatly affect a student's future in the USA, the national exam would affect a student's future in China. HS X, in part, was using the affect system to try to create a more dynamic learning environment for everyone, in the hopes of improve the overall performance on national exams from their students.

REFERENCES CITED

Bates, J, Y-W Lin, and P. Goodale. 2016. Data Journeys: Capturing the Socio-Material Constitution of Data Objects and Flows. *Big Data & Society* 3(2): DOI: 10.1177/2053951716654502.

Bennett, Jane. 2009. Vibrant Matter: A Political Ecology of Things, Durham: Duke University Press.

Besky, Sarah and Alex Blanchette. 2019. How Nature Works**:** Rethinking Labor on a Troubled Planet. Ann Arbor, MI: University of Michigan Press.

Blackman, James. 2019. Surveillance Cams to Take 70% of 5G IoT in 2020. Enterprise IOT Insights website, Oct. Accessed October 24, 2019 https://enterpriseiotinsights.com/20191024/connected-cars-2/surveillance-cams-and-c-v2x-to-take-70-of-5g-iot-share-says-gartner?utm_campaign=20191024%20Enterprise%20IoT%20NewsletterThurs&utm_medium=email&utm_source=Eloqua

Brewster, Thomas. 2019 London Police Facial Recognition 'Fails 80% Of The Time And Must Stop Now.' *Forbes* website, July 4. Accessed July 31, 2019 https://www.forbes.com/sites/thomasbrewster/2019/07/04/london-police-facial-recognition-fails-80-of-the-time-and-must-stop-now/#54fdcdf0bf95

Burrell, Jenna. 2016. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data and Society* website, Jan. Accessed May 15, 2019 https://journals.sagepub.com/doi/full/10.1177/2053951715622512

Business Times. 2019. China Shoppers Adopt Facial Recognition Payment Technology. *Business Times* website, Sept 5. Accessed Sept 5, 2019 https://www.businesstimes.com.sg/technology/china-shoppers-adopt-facial-recognition-payment-technology

Crawford, Kate and Jason Schultz. 2013. Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms. *Boston College Law Review* 55, 93-128.

Caronia, Letizia and Luigina Mortari. 2015. The Agency of Things: How Spaces and Artefacts Organize the Moral Order of an Intensive Care Unit. *Social Semiotics*, 25 (4): 401-422.

Collier, Stephen and Aihwa Ong (eds.) 2008. *Global Assemblages: Technology, Politics, and Ethics as Anthropological Problems*. Malden, MA: Blackwell.

Deluze, Gilles and Félix Guattari (Translation and Introduction by Brian Massumi). 1987. A Thousand Plateaus: Capitalism and Schizophrenia. Minneapolis: University of Minnesota Press.

Dwork, Cynthia & Deirdre K. Mulligan. 2013. It's Not Privacy, and It's Not Fair, 66 *Stanford Law Review*. 35:36–37.

Elish, Madeleine Clare. 2018. "The Stakes of Uncertainty: Developing and Integrating Machine Learning in Clinical Care." *2018 Ethnographic Praxis in Industry Conference Proceedings*, pp. 364–380, New York, NY: Wiley-Blackwell.

Einhorn, Erin. 2019. A Fight Over Facial Fecognition is Dividing Detroit — with High Stakes for Police and Privacy. *NBC News* website, August 22. Accessed October 1, 2019 https://www.nbcnews.com/news/us-news/fight-over-facial-recognition-dividing-detroit-high-stakes-police-privacy-n1045046.

Ellis, Rebecca. 2019. Portland Considers Banning Use Of Facial Recognition Software In Private Sector. *OPB News* website, Sept 17. Accessed Sept 17, 2019 https://www.opb.org/news/article/portland-facial-recognition-software-private-sector-use-ban/

Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile Police & Punish the Poor*. St : New York: Martin's Press.

Foucault, Michel (Alan Sheridan translation). 1995. *Discipline and Punish: The Birth of the Prison*. New York: Vintage Books.

Garvie, Claire. 2019. Garbage In, Garbage Out. *Georgetown Law's Center on Privacy and Technology* website, May 16. Accessed July 10, 2019 https://www.flawedfacedata.com

Gibson, William. 1999. The Science in Science Fiction. *Talk of the Nation, NPR*, November 30, Timecode 11:55. Accessed on July 9th, 2019 https://www.npr.org/2018/10/22/1067220/the-science-in-science-fiction

The Guardian. 2019. The Guardian View on Facial Recognition: a Danger to Democracy. *The Guardian* website, June 9. Accessed June 9, 2019 https://www.theguardian.com/commentisfree/2019/jun/09/the-guardian-view-on-facial-recognition-a-danger-to-democracy

Harwell, Drew. 2019. Facial-recognition Use by Federal Agencies Draws Lawmakers' Anger. *Washington Post* website, July 9th. Accessed July 9, 2019 https://www.washingtonpost.com/technology/2019/07/09/facial-recognition-use-by-federal-agencies-draws-lawmakers-anger/

Horvitz, Eric and Deirdre Mulligan. 2015. Data, Privacy, and the Greater Good. *Science*, 349 (6245):253-255.

Introna, Lucas and Helen Nissenbaum. 2009. "Facial Recognition Technology: A Survey of Policy and Implementation Issues," Report of the Center for Catastrophe Preparedness and Response, NYU.

Latour, Bruno. 2005. *Re-assembling the Social*. New York: Oxford University Press.

Kohn, Eduardo. 2013. How Forests Think: Toward an Anthropology Beyond the Human. Berkeley, CA: University of California Press.

Metz, Rachel. 2019. Amazon Wins Facial-Recognition Vote, but Concerns about the Tech Aren't Going Away. *CNN* website, May 22. Accessed July 19, 2019 https://www.cnn.com/2019/05/22/tech/amazon-facial-recognition-vote/index.html

Milgram, Stanley. 1972. "The Familiar Stranger: An Aspect of Urban Anonymity". In *The Division 8 Newsletter*, *Division of Personality and Social Psychology*. Washington: American Psychological Association

Mintz, Sidney. 1985. *Sweetness and Power: The Place of Sugar in Modern History*. New York:Viking-Penguin.

Newman, Lily Hay. 2019. Facial Recognition Has Already Reached Its Breaking Point. *Wired* website, May 22. Accessed May 29, 2019 https://www.wired.com/story/facial-recognition-regulation/

Noble, Safiya. 2016. *Algorithms of Oppression: Race, Gender and Power in the Digital Age*. New York: NYU Press.

O'Neil, Kathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York:Crown Publishing.

Pasquale, Frank. 2016. *The Black Box Society: The Secret Algorithms That Control Money and Information.* Boston: Harvard University Press.

Raji, Inioluwa Deborah & Joy Buolamwini. 2019. "Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial A.I. Products". *Conference on Artificial Intelligence, Ethics, and Society*.

Ravani, Sarah. 2019. Oakland Bans Use of Facial Recognition Technology, Citing Bias Concerns. *San Francisco Chronicle* website, July 17. Accessed July 17, 2019 https://www.sfchronicle.com/bayarea/article/Oakland-bans-use-of-facial-recognition-14101253.php

Seaver, Nick. 2017. Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems. *Big Data and Society* website, Nov. Accessed May 15, 2019 https://doi.org/10.1177/2053951717738104

Stat, Nick. 2019. "Orlando Police Once Again Ditch Amazon's Facial Recognition Software." *The Verge,* July 18. Last Accessed August 1, 2019: https://www.theverge.com/2019/7/18/20700072/amazon-rekognition-pilot-program-orlando-florida-law-enforcement-ended

Stayton, Erik, Melissa Cefkin and Jingyi Zhang. 2017. "Autonomous Individuals in Autonomous Vehicles: The Multiple Autonomies of Self-Driving Cars." 2017 *Ethnographic Praxis in Industry Conference Proceedings*, pp. blah blah New York, NY: Wiley-Blackwell.

Stayton, Erik, and Melissa Cefkin. 2018. "Designed for Care: Systems of Care and Accountability in the Work of Mobility." *2018 Ethnographic Praxis in Industry Conference Proceedings*, pp. 334–350, ISSN 1559-8918

Teicher, Jordan. 2019. What Do Facial Recognition Technologies Mean for Our Privacy? *New York Times* website July 18. Accessed July 18, 2019 https://www.nytimes.com/2018/07/18/lens/what-do-facial-recognition-technologies-mean-for-our-privacy.html

Thadani, 2019. San Francisco Bans City Use of Facial Recognition Surveillance Technology. *San Francisco Chronicle* website, May 14. Accessed July 5, 2019 https://www.sfchronicle.com/politics/article/San-Francisco-bans-city-use-of-facial-recognition-13845370.php

Tsing, Anna Lowenhaup. 2015. *The Mushroom at the End of the World: On the Possibility of Life in Capitalist Ruins*. Princeton, NJ: Princeton University Press

Trump, Donald. 2017. Executive Order 13769: Protecting the Nation from Foreign Terrorist Entry into the United States. March 6, 2017. https://www.whitehouse.gov/presidential-actions/executive-order-protecting-nation-foreign-terrorist-entry-united-states-2/

Wang, Cunrui, Qingling Zhang, Wanquan Liu, Yu Liu and Lixin Miao. 2018. Facial Feature Discovery for Ethnicity Recognition. WIREs: Data Mining and Knowledge Discovery. Wiley. July 2018. https://doi.org/10.1002/widm.1278

White, Goeff. 2019. Use of Facial Recognition Tech 'Dangerously Irresponsible' *BBC News* website, May 13. Accessed July 10, 2019 https://www.bbc.com/news/technology-48222017