Published in *Artificial Intelligence for Better or Worse* in 2019 by Future Leaders (www.futureleaders.com.au)

Big Brother walks into an office ...

Niels Wouters

A 300% increase in three years; from 200 million in 2017 to 626 million in 2020. This astounding number is the most recent prediction for the total number of surveillance cameras across China (Qiang, 2019). Close to a single camera per 2,000 inhabitants, all fully networked to form the Skynet project — perhaps an awkward name, as Skynet in *The Terminator* movies attempted to exterminate humanity (Sherman, 2017). For a long time, cameras may have been our go-to characteristic of the surveillance society, but only in recent years has the technology gained powerful technological capabilities. While camera footage would traditionally have been reviewed and annotated retroactively by human reviewers, breakthroughs in artificial intelligence (AI) and machine learning (ML) now allow for near real-time analysis by computing systems. Analysis of camera footage has become much more accessible and reliable as video feeds are scrutinised to provide us with a structured understanding of objects, changes and patterns.

The increase in application of ground-breaking AI-driven algorithms seems unstoppable and has infiltrated all layers of society. This includes health improvements through robotic surgery, financial market developments through real-time automated investment strategies, and new frontiers in sustainability through automation of the public and private transportation industry, as well as particular societal benefits, such

mental health care, as illustrated in this book. Needless to say, these developments have already impacted our daily lives and will further impact us in increasingly more profound ways.

Is AI a new tool for safer societies?

Developments of the last few years have also challenged our traditional notions of surveillance and security. Surveillance cameras and related technologies are increasingly connected to powerful AI systems that transform them into devices that predict crime and automate policing. Improvements to the efficiency of law enforcement are (for the most part) welcome developments, with comprehensive camera coverage helping us to feel and be more secure, and assisting law enforcement to better respond to security threats. While the public slowly comes to terms with the reduced privacy as these technologies become commonplace, new concerns have arisen about the use of automated decision-making algorithms in surveillance camera networks.

Cameras are no longer passive instruments that simply record particular occurrences. Instead, real-time camera feeds increasingly set in motion an opaque suite of algorithms that autonomously make decisions without any human involvement (Davenport & Harris, 2005). Imagine that driving through a red light results in a photo of your car being taken that is automatically analysed to reveal your license plate, retrieve your home address, and send the fine to your home address — all without any human involvement. Or that an in-vehicle sensor system detects a driver's drowsiness and automatically parks the vehicle in a secure location, despite contrary instructions from that driver. Or that a surveillance camera automatically (and

unmistakably) identifies a passerby as a wanted criminal and subsequently notifies the police. Applications such as these that aim to increase public safety have unmistakable benefits for society.

Prospects turn slightly grimmer when automated decision-making algorithms are not used for the benefit of society but rather for the purpose of state-sponsored surveillance or for the commercial benefit of a particular corporation. One of the most contentious government-supported systems involves China's Social Credit System, a social management and compliance program that integrates neatly with Skynet and that rewards or punishes citizens based on a computer-driven moral assessment of their everyday actions (Liang, Das, Kostyuk, & Hussain, 2018). Citizens in Western countries are perplexed by the far-reaching societal implications of this system (Botsman, 2017; Creemers, 2018). However, as other scholars point out, societies in the West have in fact already accepted a vast range of social credit and rating systems (Backer, 2017; Rosamond, 2019), ranging from Uber driver ratings and Netflix reviews to financial credit ratings and reward credit cards.

The main point of difference is that social credit in the West is largely governed by industry, offering customers relative convenience and improved service in return. The systems are not governed by a central authority and do not "speak" to each other — they do not exchange information that enables detailed user profiles to be constructed. However, the landscape is rapidly evolving. Industry increasingly experiments with new technologies that make scoring, ranking, evaluating and assessing people an intrinsic part of new

products and that provide in-depth insights for clients. Besides the so-called smart speakers, intelligent billboards are among the most established examples in this domain (McStay, 2017). In a not- too-distant future, all out-of-home advertising may be curated by AI algorithms that completely personalise the types of products offered and their prices. Assumptions made by such AI algorithms may be based on objective, visual characteristics of the viewer (e.g. based on your clothing, apparel and make-up). Assumptions made by such AI algorithms may also include the characteristics of the people around you — and not necessarily limited to people physically proximate to you, but also your Facebook, Twitter or Instagram friend networks. This may all be fine if the algorithm's assumptions reflect your true spending power, but it quickly turns problematic if the algorithm incorrectly identifies you as a millionaire.

These scenarios are not far-fetched and actually encourage other industries to start experimenting in this area. Would it be similarly problematic if your insurers' algorithms analyses road safety camera footage and adjusts your insurance premium in response to your driving style? Would it be problematic if a recruiter uses an algorithm to analyse video footage of your job interview to assess suitability for a job? Whilst some of us may not necessarily oppose these particular approaches, the challenge remains to draw the line between what is technologically possible and what is societally acceptable (Wouters et al., 2019).

What Could Workspaces of the Future Feel Like?

Marcia Hayes, a customer service specialist for a digital service provider, enters the office building

every day by walking through high-tech security gates. No need for access badges anymore as cameras compare a live photo of Marcia's face with a three-dimensional facial scan in the company database. Besides her time of arrival, other metrics that are captured and annotated include pupil dilation, breath, skin tone, gait and a brief voice sample. After working for the company for 7 years, it's a routine that Marcia is now very familiar with. Today is no different, but she briefly questions whether she should have drunk that second glass of wine last night — it seems to take the security gate slightly longer than usual to process her arrival. Her walk to the elevator and from the elevator to her desk are timed by correlating facial recognition markers on the company's vast CCTV network.

As Marcia sits down, her PC monitor turns on and Sam — the company's in-house AI — prompts Marcia that her work hours have officially begun. A subsequent prompt notifies Marcia that she arrived 27 seconds later than yesterday. This is unfortunate, as arrival times are gamified in the company: she drops from the 15th to the 27th place and is ruled ineligible for company-sponsored Friday night drinks for the 20 best performing employees. Marcia lets out a big sigh; she realises that she should have tied shoelaces before walking through the security gates on the ground floor (not after). As she closes the notification, her PC monitor automatically increases brightness in response to Marcia's current pupil dilation, and the first customer phone call comes through.

Hours later Sam prompts Marcia to take a break. She's been on calls for three hours straight now and her voice is starting to sound fatigued. Sam knows that green tea would be most suitable for Marcia at this stage, based on her current state and the morning metrics. Conveniently, Sam offers to automatically order a fresh brew from the office barista — Marcia will be informed when it's ready for pickup. Marcia gladly accepts the offer, knowing that accepting Sam's proposals benefits her performance score. The cost of the tea is automatically deducted from Marcia's pay.

Later that afternoon Sam informs Marcia that she'll have to stay back — an animated emoticon accompanies the automated message. Seemingly, Marcia did not perform well enough in 13% of her customer phone calls today. While no extra information as to Marcia's subpar performance are given, Sam knows that 47 minutes of overtime are sufficient. Marcia has no choice but to accept the decision, and nervously looks around the office floor as she knows this information also appears on the large LED screen that serves as the team's leader board. Perhaps, Marcia thinks for a brief second, it's time to look for other job opportunities. Sam detects Marcia's disengagement and prompts her to stay focussed. Another customer phone call comes through.

Is AI a new tool for better performing employees?

The office experience of Marcia is far-fetched and influenced heavily by my interest in science fiction. However, some of what Marcia encounters is real and already operating in corpo-

rate call centres, in transit hubs, and among sports players and couriers (Shell, 2018; Simonite, 2018). At some point it is likely to make its way into your office too.

Algorithm-driven surveillance and assessments are increasingly finding their way into the workspace, integrated within so-called workplace surveillance tools.[1] Obvious opportunities exist in automating repetitive tasks by way of advanced technologies and conducting intelligent and unobtrusive error checks to help mitigate financial risks, privacy breaches or other harm to the business. However, more and more there is a tendency towards connecting workplace surveillance cameras — yes, what was formerly considered to be spy technology — and device usage data to advanced AI. Connecting these functions allows monitoring of the activities and behaviours of employees on the workfloor by methods including audio and video analytics, and behaviour and activity recognition. Surveillance is no longer discrete and limited to supervisors who sporadically cast their eye across the office, but has rather become an omnipresent and continuous monitoring system. The main purpose of surveillance has long been to evaluate employee performance, health and temperament, and their movements. However, recent developments enable employers to gather complex data in real time and to immediately inform strategic decision making. In fact, recent developments even enable employee monitoring far beyond the physical boundaries of the workplace. Most current practical applications range from logging typed text on keyboards and documenting opened and printed documents, to recording conversations with customers and between coworkers, tracking movement throughout the office building by way of cameras and biomet-

ric authentication systems, and tracking employee health data by way of connected wearable devices that employees may be required to carry while at work.

The general aim of workplace surveillance is obvious: to analyse business operations in order to increase productivity and profitability. Have employees reached a prescribed (but perhaps ill-defined) number of keystrokes, and who is underperforming? Are employees spending an excessive (but perhaps ill-defined) amount of time near the coffee machine? Are conversations between employees about important and appropriate (but perhaps ill-defined) topics? These developments are not hard to fathom from a commercial point of view, as employees' activities are instrumental to the success of any business. However, as more and more datapoints become available, managers have gained access to a digital panopticon. It allows them to oversee coworkers in real time and to adjust daily operations of the business in response to an ever-growing suite of metrics.

We can only assume that these metrics are mere modern takes on traditional approaches, such as understanding an individual's success to generate new business leads or measuring the time needed to complete a routine task. However, the wealth of new digital tools available now enables employers to revive the principles of early 20th century Taylorism (Taylor, 1919), a reductionist approach that dehumanises the employee by solely focusing on maximising their efficiency in machinic and mechanistic ways. The same developments that help to make society more secure are now harnessed as a weapon against our own workforce through a new form of postmodern surveillance (Holford, 2019). Drug and DNA tests

may be fairly established in the human resources process, but also workers' real-time performance can now be analysed by AI-driven processes at an alarmingly high rate. Automated decision-making processes can then be set in motion in the event that employee behaviours and metrics fall behind or exceed predefined boundaries. These processes are riddled with reasons for concern.

First and foremost, management is likely to deny the existence and usage of surveillance algorithms in the workplace, as there is no legal obligation in most regions to be fully transparent. Although highly concerning for workers, the absence of legal frameworks means that management can quietly roll out its own, privately operated surveillance networks. Management can use the secretly obtained information to influence its decision making, while suppressing workers' critiques and avoiding efforts to avoid or manipulate the surveillance systems. We must also assume that management can remain vague about the specific metrics that define good and bad performance. Unlike quantitative KPIs or even Objectives and Key Results (OKRs), these metrics focus on qualitative, personal and behavioural characteristics that are prone to subjective interpretation — whether by a person or an algorithm.

Perhaps more important is that the secrecy prevents resistance from employees and criticism from the public, who would happily draw parallels with dystopian classics such as *Brazil, 1984* and *Black Mirror*, and boycott further dealings with the business. This leads to an additional concern, as employees are in fact denied a basic right: transparency and ownership over the data concerning their individual perform-

ance and behaviours — they may only experience the final decision: hired, rejected, transferred, promoted or sacked. Absence of transparency prevents employees from retaining full oversight over the factors that ultimately influence career progress. Beside employees, we must assume that human resources staff also remain unaware of the data and decision-making factors that affect workers' careers. Lack of oversight into the entire dataset that leads to an (automated) decision can leave human resources staff bewildered about the reasons that underpin hiring and dismissal decisions. Lack of oversight risks turning the human resources processes and staff management into fully automated activities that are devoid of any humanness and that fail to build employees' confidence over their career.

The root of the problem lies with current technology providers. Hiding behind claims of intellectual property, they avoid public insight into the inner workings of the algorithms they develop and resell. Out-of-the-box, well-designed dashboards for live employee monitoring systems are typically the only components buyers will get to see (Lewis, 2017) — the data processing and decision making instead occurring on distributed cloud-connected servers that operate as black box systems, opaquing all inner workings to anyone but the technology provider itself (Pasquale, 2015). In fact, as AI and ML algorithms are increasingly training and retraining themselves based on the data that feeds into them (dubbed "deep reinforcement learning"; Sutton & Barto, 2018), understanding the inner workings of black box systems (even to the technology providers themselves) becomes more and more challenging. A recent and highly publicised case that illustrates

the social concerns of black box systems involved multinational technology corporation Amazon (Fernandez, 2018). It developed an AI-assisted hiring tool in-house that organically developed a bias against women due to implicit biases in its training data, that is, an archive of resumes previously submitted for male-dominated technical roles. The bias was revealed only after a thorough review was conducted by humans and the project was subsequently suspended.

However, part of the problem lies with employees themselves. For instance, who of you does not send a personal email via a work computer, be it in the office or while working from home? Or who doesn't log on to a personal email service on that same work computer? It's fair to say that we all do. But in doing so, we explicitly authorise employers to monitor, check, log and otherwise scrutinise the contents of messages, browser behaviour and the nature of any other interaction with an electronic communication service.[2] Why? Most likely because the computer is their property, or because we use our employer's network to access these services. Common sense urges employers to prepare internet and email usage policies and educate staff accordingly. However, as boundaries between private and professional technology usage blur, employees must also develop a healthy understanding of privacy-conscious technology usage and their potential professional, personal and social ramifications. This is explained in great detail in Solomon and Andersen's chapter in this book.

How do we proceed from here?

We have the potential to introduce technology in the workplace as a means to promote business performance, by

stimulating customer satisfaction and retention as employees are encouraged to thrive in productive, interactive and welcoming environments. But, if we fail to do so, we are at risk of reaffirming the machinic and purely efficiency-focused approaches that were established during the Industrial Revolution. As the title of this chapter suggests, Big Brother increasingly permeates our office environments. While the title seems to suggest a satirical take on the issue, workplace surveillance is, in fact, an urgent issue that must be addressed at the social, political, legal and technological level.

A lot of work remains to be done in the realm of automated decision making in the context of work. If an algorithm were to rule Marcia unfit for the role because of a range of qualitative and quantitative metrics, these factors should be communicated transparently and explained unambiguously to her. Were the 13% of today's phone calls where she underperformed instrumental to the decision, or is there more? How was the decision made? Which are the targets that employees at the firm should work towards? But, also, who owns these data while Marcia works for her current employer, but also when she is employed elsewhere?

Now is the time to pause further deployments of spy technology in the office and reconsider the objectives behind their installation and usage. Instead, let's take a break and formulate strategies that warrant ethical, meaningful and professionally encouraging applications. Beyond the legal requirements that are currently in place (which, in Australia is limited to NSW,[3] ACT[4] and to some extent VIC[5]), let's empower employees by providing full disclosure about the presence and operation of intelligent workplace surveillance systems. Full

disclosure would at least imply that employees are made aware of the motivations behind the installation of these systems, their operational parameters and constraints, the people that are authorised to review footage, records and analyses, and — importantly — the metrics and policies that govern the implications of surveillance analysis. Let's assume that the analysis is used for more than policing and punishment, but instead encourages staff development, career progression and healthy work environments.

As a result, if workplace monitoring is an unstoppable trend, measures must be put in place to guarantee that monitoring happens in transparent ways that allow for review and discussion with the stakeholders that are involved. Enabling managers to view staff analyses in easy-to-use dashboards suggests an opportunity for staff members to review their own, personal data on those same platforms. There is a strong case for explainability, as explained in Tim Miller's chapter in this book, to become an intrinsic part of workplace surveillance. Only through these extended forms of transparency can data-driven career discussions become democratic and mutually beneficial. In fact, it may even help broaden the discussion about the metrics, parameters and values that we consider useful to be monitored in the workplace.

And ultimately the discussion boils down to thousands of datapoints on server infrastructure. We must recognise that datapoints tend to exist forever, even well after their owners have moved on. As rating and review systems increasingly interconnect, here is an enormous opportunity to alleviate public concern: as employees move on, change careers and make life-changing decisions, we should put mechanisms in

place that protect their wellbeing. Performance data generated in the workplace should not live on forever so that they can be revisited as career paths have changed. Many factors may have contributed to one's performance at an employer, for better or worse. Family situations change, career interests evolve, our social networks change. I can only hope that we are not progressing towards a society where all of these parameters are given the consideration they deserve as we consider augmenting the workplace with surveillance, monitoring and decision-making systems.

We are at a particularly interesting point in time where we have an enormous suite of technological tools available that affect our lives and careers. However, at the same time, there is a growing challenge to use technology for the benefit of society rather than to its detriment. As more datapoints become available, the potential for misuse and misinterpretation grows. And as deployments in foreign nations show, surveillance technology is particularly prone to misuse. Now is the time to start a broad and inclusive discussion about where we want these technologies in the office to take us, what parameters we put in place to prevent unfreedom, and how we can use data to encourage employee satisfaction. With this chapter, I hope to make tangible the exact challenges that we are facing, in the hopes that this facilitates an accessible, inclusive and useful discussion about the future of work.

Endnotes

1 Workplace surveillance typically consists of methods such as camera surveillance (where visual images of activities and behaviors are recorded), computer surveillance (where information input or output is monitored or recorded) and tracking surveillance (where

location of movement are tracked). To best illustrate the general implications of workplace surveillance, I do not differentiate between any of those methods in this chapter.

2 A great resource for both employees and employers is the Australian Government's Fair Work Ombudsman Best Practice guide on Workplace Privacy. https://www.fairwork.gov.au/how-we-will-help/templates-and-guides/best-practice-guides/workplace-privacy#emailinternet

3 *Workplace Surveillance Act 2005* No 47.

4 *Workplace Privacy Act 2011*.

5 *Surveillance Devices Act 1999*.

References

Backer LC (2017). *Measurement, assessment and reward: The challenges of building institutionalized social credit and rating systems in China and in the West*. Paper presented at the The Chinese Social Credit System, Shanghai Jaiotong University.

Botsman R (2017, October 21). Big Data Meets Big Brother as China moves to rate its citizens. *Wired*. https://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion

Creemers R (2018). China's Social Credit System: An evolving practice of control. *SSRN Electronic Journal*. http://doi.org/10.2139/ssrn.3175792

Davenport TH & Harris JG (2005). Automated Decision Making Comes of Age. *MIT Sloan Management Review, 46*, 83–89.

Fernandez J (2018). The ball of wax we call HR Analytics. *Strategic HR Review, 18*, 21–25.

Holford W-D (2019). The future of human creative knowledge work within the digital economy. *Futures, 105,* 143–154.

Lewis R (2017). *Under Surveillance: Being Watched in Modern America*. University of Texas Press.

Liang F et al. (2018). Constructing a data-driven society: China's Social Credit System as a state surveillance infrastructure. *Policy & Internet, 10*, 415–453.

McStay A (2017). Empathic media: Emotiveillance and the future of out of home advertising. In *Digital Advertising* (pp. 150–163). Macmillan Education UK.

Pasquale F (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.

Qiang X (2019). The road to digital unfreedom: President Xi's surveillance state. *Journal of Democracy, 30*, 53–67.

Rosamond E (2019). From reputation capital to reputation warfare: Online ratings, trolling, and the logic of volatility. *Theory, Culture & Society.*

Shell ER (2018, October 15). The employer-surveillance state. *The Atlantic.* https://www. theatlantic. com/business/archive/2018/10/employee-surveillance/568159/

Sherman FA (2017). *Now and Then We Time Travel: Visiting Pasts and Futures in Film and Television.* McFarland.

Simonite T (2018). This call may be monitored for tone and emotion. *Wired.* https://www.wired. com/story/this-call-may-be-monitored-for-tone-and-emotion/

Sutton RS & Barto, AG (2018). *Reinforcement Learning: An Introduction.* MIT Press.

Taylor FW (1919). *The Principles of Scientific Management.* Harper & Brothers.

Weston M (2015). Wearable surveillance — A step too far? *Strategic HR Review, 14*, 214–219.

Wouters N et al. (2019). *Biometric mirror: Exploring values and attitudes towards facial analysis and automated decision-making.* Paper presented at the Conference on Designing Interactive Systems. http://doi.org/10. 1145/3322276.3322304