OSH AND THE FUTURE OF WORK: BENEFITS AND RISKS OF ARTIFICIAL INTELLIGENCE TOOLS IN WORKPLACES

Introduction

Artificial intelligence (AI) was born, or at least its name was, in 1956, at a series of academic workshops organised at Dartmouth College in New Hampshire, United States. At that conference a group of scientists set out to teach machines to use language, form concepts, improve themselves (as machines) and solve problems originally 'reserved for humans' (McCarthy et al., 1955). John McCarthy and his colleagues had high hopes that they could achieve this within a few weeks. The conference was not successful on its own terms, but, nevertheless, a significant field of research and development in AI was launched.

We might now laugh at this optimism, but interest in AI did not disappear. Indeed, AI debates and experimentation have gone through a series of phases, from the peak of hope that machines could be trained to behave exactly like people and achieve an equivalent level of intelligence to humans, as seen at the Dartmouth workshops, to the troughs of disillusionment. The first experimental robots, such as 'WOBOT' and 'Shakey', did not achieve the universal AI they were aiming for. Two so-called AI winters lasted from 1974 to 1980 and from 1987 to1993 as various experiments failed and funding waned. But now, in 2019, a revived interest is bubbling.

Nowadays, advanced countries are allocating significant pots of funding, in the order of billions, to research and development in AI, with the United States in the lead, closely followed by China and Israel (Delponte, 2018). AI is predicted to provide a 26 % boost to gross domestic product (GDP) by 2030 in China. North America is predicted to see a 14.5 % boost (PwC, 2018a), and some predictions indicate that AI will create as many jobs as it eliminates (PwC, 2018b). Consultancies' and thinktanks' forecasts run alongside a series of governmental, regional and international organisations' high-level reports that predict the significant impact of AI on economies and societies, including the United States (White House Office of Science and Technology Policy, 2018); the United Kingdom's Department for Business, Energy and Industrial Strategy and Department for Digital, Culture, Media and Sport (2018); the International Labour Organization (ILO) (Ernst, Merola and Samaan, 2018); and the European Union (European Commission, 2018).

In most cases, high-level governmental and organisational reports are predicting that AI will improve productivity. Discussions of productivity involve direct implications for workers and working conditions, of course, but there is little discussion yet of how the introduction of AI into workplaces will benefit or put at risk occupational safety and health (OSH) for workers themselves. To lay the foundations for this expert report, which addresses this gap, we start by discussing the meaning of AI to provide a clear steer for discussions of its impact on workers. Then, in Chapter 2, we outline where AI is being used in a series of applications and tools for assisted work, as well as workplace decision-making, and the OSH risks and benefits arising. We start with human resources (HR) via people analytics and interview filming, then look at the integration of AI-augmented robotics, including collaborative robots (cobots) and chatbots in factories, warehouses and call centres. Next, we identify uses of wearable technologies and assistive tablets on the production assembly line and then outline algorithmic processes in work in the gig economy. Then, in Chapter 3, we outline international stakeholder responses to the rising risks and benefits of AI at work. In conclusion, in Chapter 4, the report provides some recommendations for how to best manage and mitigate the worst risks that could arise from using AI in workplaces.

1 What is AI?

There is debate today about 'what is AI' and 'what is not AI'. It may even appear that there is more hype around AI than reality. Nonetheless, as governments are pouring huge amounts of capital into research and development and publishing high-level reports making notable predictions about

the contributions that AI will make to GDP and productivity, it is worth taking AI seriously. The dispute around the authenticity of AI is relevant, however. So, rather than waver on the definition throughout this report, the original discussion about what AI 'could be' is recalled. McCarthy and his colleagues, mentioned in the introduction, defined the 'artificial intelligence problem' as one that 'is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving' (McCarthy et al., 1955). Since the authors of the Dartmouth document invented the concept of AI, recalling their definition lends much to the discussion. Can machines behave like humans? This philosophical question is not extensively dealt with in this article, but it is worth noting that wider questions about humans and our relationship with machines were central to this research area's early incarnations (see, for example, Simon, 1969; Dreyfus, 1972; Weizenbaum, 1976), and they still operate in the background of AI experimentation and application today. Central to these questions is the fairly obvious, but rarely vocalised, question: *why* do we want machines to behave like us and even better than us? Socially, what is missing that we need such improvements? In any case, while there are a number of definitions of AI, for the purposes of this report, McCarthy's definition will be used as a general insight to locate the emerging issues epistemologically.

The European Commission's definition, as provided in its 2018 Communication, is adopted for this report, whereby AI 'refers to systems that display intelligent behaviour by analysing their environment and taking actions — with some degree of autonomy — to achieve specific goals' (European Commission, 2018). Another 2018 report entitled *European artificial intelligence leadership, the path for an integrated vision* further defines AI as a 'cover term for techniques associated with data analysis and pattern recognition' (Delponte, 2018, p. 11). That report, which was requested by the European Parliament's Committee on Industry, Research and Energy, differentiates AI from other digital technologies in that 'AI are set to learn from their environments in order to take autonomous decisions' (Delponte, 2018, p. 11). These definitions facilitate a clear discussion about what is at stake as AI systems and machines are integrated into workplaces, where systems demonstrate competences that allow decision-making and prediction much faster and more accurately than humans and provide human-like behaviour and assistance for workers.

There are various levels of AI now discussed by experts: weak and strong. 'Weak AI' is where a machine relies on software to guide its investigation and responses. This type of AI does not reach a level of consciousness or full sentience as such, but it acts as a problem-solver within a specific field of application. 'Weak AI' thus applies to expert systems and text and image recognition. 'Strong AI', also called 'universal AI' (Hutter, 2012), on the other hand, refers to when a machine can demonstrate behaviour that equals or exceeds the competence and skill of humans, and this is the type of AI that most intrigued researchers such as Alan Turing. Even before McCarthy and his colleagues' conference in 1956, in 1950, Alan Turing had asked himself, 'Can machines think?' (Turing, 1950). The stage of universal AI is reached when a single universal agent can learn to behave optimally in any environment, where universal competences are demonstrated by a robot, such as walking, seeing and talking. Today, as computer memory capacity increases and programmes become more sophisticated, universal AI is becoming increasingly likely. This is an advance that could complete the automation process, whereby robots become as good at working as people and do not exemplify human characteristics such as tiredness or sickness, and so on. People appear to feel more comfortable with weak AI, which enhances machines and means that they behave like assistants to humans, rather than replacing us as workers or replacing human management.

We now outline the uses of AI at work and the potential and evidence for risks and benefits for OSH, based on desk-based research and a series of expert interviews carried out by the author.

2 AI in the workplace

Although there are significant possibilities for workplace progress and growth in productivity, there are also important OSH safety and health-related questions arising as AI is integrated into workplaces. Stress, discrimination, heightened precariousness, musculoskeletal disorders, and the possibilities of work intensification and job losses have already been shown to pose psychosocial risks, including physical violence in digitalised workplaces (Moore, 2018a). These

risks are exacerbated when AI augments already existing technological tools or are newly introduced for workplace management and design. Indeed, AI exaggerates OSH risks in digitalised workplaces, because it can allow increased monitoring and tracking and thus may lead to micro-management, which a prime cause of stress and anxiety (Moore, 2018a). AI stresses the imperative of giving more credibility and potentially authority to what Agarwal and colleagues (2018) call 'prediction machines', robotics and algorithmic processes at work. But it is worth stressing that it is not technology in isolation that creates OSH benefits or risks. It is instead the *implementation* of technologies that creates negative or positive conditions.

2.1 AI in human resources

In the area of HR business execution, one increasingly popular area of AI integration is called 'people analytics', defined broadly as the use of big data and digital tools to 'measure, report and understand employee performance, aspects of workforce planning, talent management and operational management' (Collins, Fineman and Tsuchida, 2017). Computerisation, data gathering and monitoring tools allow organisations to conduct 'real-time analytics at the point of need in the business process ... [and allow] for a deeper understanding of issues and actionable insights for the business' (ibid.). The prediction machine algorithms applied for these processes often reside in a 'black box' (Pasquale, 2015), and people do not fully understand how they work, but, even so, computer programs are given the authority to make 'prediction[s] by exception' (1) (Agarwal, Gans and Goldfarb, 2018).

Not all people analytics have to be, strictly speaking, AI. However, programs' intelligent responses to algorithmic equations allow machine learning, which generates predictions and asks associated questions that emerge without human intervention except at the data input phase, and are AI in the sense of the EU's definition above. Big data has been seen as a lucrative growth area for some years, whereby the collection of information about everything, all the time, has been an attractive investment. Now, the big data era is paying off in HR circles, because the extensive pools of data now available can be used to train algorithms to form analyses and make predictions about workers' behaviour via machine learning and thereby assist management decision-making. On the basis of the patterns identified, AI enables an algorithm to produce solutions and responses to enquiries about patterns across data much more quickly than people could. Machine learning responses are often unlike those that a human alone would, or perhaps even could, generate. Data about workers can be gathered from various sources both in and outside the workplace, such as number of keyboard clicks, information from social media, number of and content of telephone calls, websites visited, physical presence, locations visited outside the workplace through GPS (global positioning system) tracking, movements around the office, content of emails and even tone of voice and bodily movements in sociometrics (Moore, 2018a, 2018b).

Also called 'human analytics',' talent analytics' and 'human resource analytics', in an era of 'strategic HR', this application of AI-enabled tools is defined broadly as the use of individualised data about people to help management and HR professionals make decisions about recruitment, i.e. who to hire, for performance appraisals and promotion considerations, to identify when people are likely to leave their jobs and to select future leaders. People analytics are also used to look for patterns across workers' data, which can help to spot trends in attendance, staff morale and health issues at the organisational level.

About 40 % of HR functions in international companies are now using AI applications. These companies are mostly based in the United States, but some European and Asian organisations are also coming on board. A PwC survey shows that more and more global businesses are beginning to see the value of AI in supporting workforce management (PwC, 2018a). One report shows that 32 % of personnel departments in tech companies and others are redesigning organisations with the help of AI to optimise 'for adaptability and learning to best integrate the insights garnered from employee feedback and technology' (Kar, 2018) Recent IBM research indicates that, in the world's 10 largest economies, even as many as 120 million workers may need

(1)'Prediction by exception' refers to processes whereby computers deal with large data sets and are able to make reliable predictions based on routine and regular data and also to spot outliers and even send notifications 'telling' the user that checks should be made or that human assistance or intervention should be provided.

to be retrained and reskilled to deal with AI and intelligent automation. This report indicates that two thirds of CEOs believe AI will drive value in HR (IBM, 2018). A Deloitte report shows that 71 % of international companies consider people analytics a high priority for their organisations (Collins, Fineman and Tsuchida, 2017), because it should allow organisations to not only provide good business insights but also deal with what has been called the 'people problem' (ibid.).

'People problems' are also called 'people risks' (Houghton and Green, 2018), which are divided into seven dimensions in a Chartered Institute for Personnel Development (CIPD) report (Houghton and Green, 2018) as:

1. talent management,
2. health and safety,
3. employee ethics,
4. diversity and equality,
5. employee relations,
6. business continuity, and
7. reputational risk.

But perhaps people are not the only 'problem'. Based on the original definition of AI, in which machines are predicted to eventually have the capability of behaving as a human would, if humans are discriminating and biased, then we should not be surprised when AI provides biased answers. In other words, machine learning operates only on the data that it is fed, and if those data reveal past discriminatory hiring and firing practices, then the results of the algorithmic process are likely to also be discriminatory. If the information gathered about workers is not buffered with qualitative information about individuals' life experiences and consultation with workers, unfair judgements could be made (see below for more on this).

AI-enhanced HR practices can help managers obtain seemingly objective wisdom about people even before they hire them, as long as management has access to data about prospective workers, which has significant implications for tailoring worker protection and preventing OSH risks at the individual level. Ideally, people analytics tools can aid employers to 'measure, report and understand employee performance, aspects of workforce planning, talent management and operational management' (Collins, Fineman and Tsuchida, 2017). Indeed, algorithmic decision-making in people analytics could be used to support workforces by aligning employee performance feedback and performance pay — and workforce costs — with business strategy and support for specific workers (Aral et al., 2012, cited in Houghton and Green, 2018, p. 5). Workers should be personally empowered by having access to new forms of data that help them to identify areas for improvement, that stimulate personal development and that achieve higher engagement.

However, if processes of algorithmic decision-making in people analytics do not involve human intervention and ethical consideration, this human resource tool could expose workers to heightened structural, physical and psychosocial risks and stress. How can workers be sure that decisions are being made fairly, accurately and honestly, if they do not have access to the data that their employer holds and uses? OSH risks of stress and anxiety arise if workers feel that decisions are being made based on numbers and data that they have neither access to nor power over. This is particularly worrying if people analytics data leads to workplace restructuring, job replacement, job description changes and the like. People analytics are likely to increase workers' stress if data are used in appraisals and performance management without due diligence in process and implementation, leading to questions about micro-management and workers feeling 'spied on'. If workers know that their data are being read for talent spotting or for deciding possible layoffs, they may feel pressurised into improving their performance and begin to overwork, posing OSH risks. Another risk arises with liability, in which companies' claims about predictive capacities may later be queried for accuracy or personnel departments held accountable for discrimination.

One worker liaison expert ([2]) indicated that worker data collection for decision-making such as that seen in people analytics has created the most urgent issues arising with AI in workplaces. Often, works councils are not aware of the possible uses of such management tools. Or, systems are

([2])Dr Michael Bretschneider-Hagemes, Head of the Employees Liaison Office of the German Commission KAN spoke to the author of this report in an interview on 18 September 2018.

being put into place without consulting with works councils and workers. Even more OSH risks arise, such as worker stress and job losses, when technologies are implemented in haste and without appropriate consultation and training or communication. In this context, it is interesting to consider a project run at the headquarters of IG Metall, in which workplace training curricula are being reviewed in 2019 in the context of *Industrie 4.0.* (see also section 3.4) ([3]). The findings demonstrate that training needs updating not only to prepare workers for physical risks, as has been standard in heavy industry OSH training, but also for the mental and psychosocial risks introduced by digitalisation at work, which includes people analytics applications ([4]).

Another form of people analytics involves filming job interviews. This practice is carried out by organisations such as Nike, Unilever and Atlantic Public Schools. These companies are using products that allow employers to interview candidates on camera, in which AI is used to judge both verbal and non-verbal cues. One such product is made by a group called HireVue and is used by over 600 companies. The aim is to reduce bias that can arise if, for example, an interviewee's energy levels are low or if the hiring manager has more affinity for an interviewee based on similar, for example, age, race and related demographics. However, there is evidence that preferences from previous hiring managers are reflected in hiring, and heterosexual white men are, a report by Business Insider reveals, the hiring preference, other things being equal (Feloni, 2017). If the data provided to an algorithm reflect the dominant bias over time, then it may score someone with 'in group' facial expressions higher and give a lower rating to other cues tied to sexual orientation, age and gender that do not resemble a white male.

Overall, people analytics poses both benefits and risks for OSH. As this tool uses algorithms, machines should be subject to extensive testing before they are used for any of the HR applications outlined. Another possibility is for a people analytics algorithm to be designed specifically to eliminate biases, which is not an easy task. Risk assessments are already being experimented with in criminal systems in which AI informs sentencing and parole boards to attempt to eliminate bias. IBM has recently publicised a tool that likewise intends to reduce risks of discrimination. It is hoped that these types of initiatives will deal with rising risks for OSH in AI-assisted HR decision-making. Nonetheless, AI's strength is also its weakness.

2.2 Cobots in factories and warehouses

We can picture the scene: huge orange robot arms in factories whirring away in expansive warehouses in industrial landscapes, building car parts and assembling cars where conveyor belts lined with humans once stood. Robots have directly replaced workers on the assembly line in factories in many cases, and sometimes AI is confused with automation. Automation in its pure sense involves, for example, the explicit replacement of a human's arm with a robot arm. EU-OSHA's report *Foresight on new and emerging occupational safety and health risks associated with digitalisation by 2025* (EU-OSHA, 2018, p. 89) indicates that robots allow people to be removed from dangerous physical work and environments with chemical and ergonomic hazards, thus reducing OSH risks for workers.

Lower skilled, manual work has historically been most at risk and is still at a high risk of automation. Now, automation can be augmented with autonomous machine behaviour or 'thinking'. So, the AI dimension of automation reflects where workers' brains, as well as their limbs, may no longer be needed. Now, as one EU-OSHA discussion paper on the future of work regarding robots and work indicates, while robots were at first built to carry out simple tasks, they are increasingly enhanced with AI capabilities and are being 'built to think, using AI' (Kaivo-oja, 2015).

([3]) *Industrie 4.0* is a much-debated term that originated in German manufacturing circles, designed to advance manufacturing in marketing terms. Some critics argue that it is a narrative rather than a reality today. Nonetheless, it is commonly accepted that, if there is to be a trajectory of industrial revolutions, *Industrie 1* is the term for the first industrial revolution and thus the invention of the steam engine. The second is linked to advancements in science and the third to digitalised inventions incorporated into production. Today, the 'Internet of Things', whereby machines technically communicate with another, advanced robotics and increased capacity for memory and processing power are seen as the driving force for the concept of *Industrie 4.0*.

([4]) Antje Utecht, who works in the training and policy department at the headquarters of IG Metall in Frankfurt, Germany, shared these insights with the author of this report during an interview on 16 October 2018.

Cobots are now being integrated into factories and warehouses where they work alongside people in a collaborative way. They assist with an increasing range of tasks, rather than necessarily automating entire jobs. Amazon has 100,000 AI-augmented cobots, which has shortened the time taken to train workers to less than 2 days. Airbus and Nissan are using cobots to speed up production and increase efficiency.

As a recent Netherlands Organisation for Applied Scientific Research (TNO) report states, there are three types of OSH risks in human-cobot-environment interactions (TNO, 2018, pp. 18-19):

1. robot-human collision risks, in which machine learning can lead to unpredictable robot behaviour;
2. security risks, in which robots' internet links can affect the integrity of software programming, leading to vulnerabilities in security; and
3. environmental risks, in which sensor degradation and unexpected human action in unstructured environments can lead to risks to the environment.

AI-permitted pattern and voice recognition and machine vision mean that not only are unskilled jobs at risk of replacement but now a range of non-routine and non-repetitive jobs can be carried out by cobots and other applications and tools. In that light, AI-enhanced automation enables many more aspects of work to be done by computers and other machines (Frey and Osborne, 2013). One example of the protection of workplace OSH through AI-augmented tools is found in a chemicals company that makes optical parts for machines. The miniscule chips that are produced need to be scanned for mistakes. Previously, one person's job was to detect mistakes with their own eyes, sitting, immobile, in front of repeated images of chips for several hours at a time. Now, AI has fully replaced this task. The OSH risks, which have now been of course, eliminated, include musculoskeletal disorders and eye strain and damage ([5]).

Cobots can reduce OSH risks, as they allow AI systems to carry out other types of mundane and routine service tasks in factories, which historically create stress, overwork, musculoskeletal difficulties and even boredom as a result of repetitive work. However, AI-augmented robots in factories and warehouses can create stress and a range of serious problems if they are not implemented appropriately. Indeed, one UK-based trade unionist indicated that digitalisation, automation and algorithmic management when 'used in combination … are toxic and are designed to strip millions of folks of basic rights' ([6]). Potential OSH issues may also include psychosocial risk factors if people are driven to work at a cobot's pace (rather than the cobot working at a person's pace) and collisions between cobots and people ([7]). Another cobot-related case of machine-human interaction creating new working conditions and OSH risks is when one person is assigned to 'look after' one machine and is sent notifications and status updates about machines on a personal device such as a smartphone or a home laptop. This can lead to risks of overwork, whereby workers feel compelled to take note of notifications in out-of-work hours and their work-life balance is disrupted ([8]).

One expert ([9]) in AI and work discussed developments around the Internet of Things in workplaces, in which machine-to-machine connected systems work alongside human labour in factories and warehouses. Data input problems, inaccuracies and faults with machine-to-machine systems create significant OSH risks as well as questions of liability. Indeed, sensors, software and connectivity can be faulty and unstable, and all vulnerabilities raise questions about who is legally responsible for any damage that emerges. Is it a cobot's fault if it runs into a worker, the worker's fault, the company who manufactured the cobot originally or the company that employs the worker and integrates the cobot? The complexities abound.

Human-robot interaction creates both OSH risks and benefits in the physical, cognitive and social realms, but cobots may someday have the competences to reason and therefore must make humans feel safe. To achieve this, cobots must demonstrate perception of objects versus humans

([5]) Information obtained from interview with Antje Utecht (fn 4).
([6]) Interview with Maggie Dewhurst of the Independent Workers Union of Great Britain (IWGB) in 2017.
([7]) Based on an interview with Dr Sam Bradbrook, specialist at the United Kingdom's Health and SafetyExecutive'sForesightCentre, in September 2018.
([8]) Interview with Antje Utecht (fn 4).
([9]) Interview with Dr Sam Bradbrook (fn 7).

and the ability to predict collisions, adapt behaviour appropriately and demonstrate sufficient memory to facilitate machine learning and decision-making autonomy (TNO 2018, p. 16) along the lines of the previously explained definitions of AI.

2.3 Chatbots in call centres

Chatbots are another AI-enhanced tool that can deal with a high percentage of basic customer service queries, freeing up humans working in call centres to deal with more complex questions. Chatbots work alongside people, although not only in the physical sense. Within the back end of systems, they are used to deal with customer queries over the phone using natural language processing. Dixons Carphone uses a conversational chatbot now named Cami that can respond to first-level consumer questions on the Currys website and through Facebook Messenger. Insurance company Nuance launched a chatbot named Nina to respond to questions and access documentation in 2017. Morgan Stanley has provided 16,000 financial advisers with machine learning algorithms to automate routine tasks.

Call centre workers already face extensive OSH risks because of the nature of the work, which is repetitive and demanding and subject to high rates of micro-surveillance and extreme forms of measurement (Woodcock, 2016). An increasing number of activities are already recorded and measured in call centres. Words used in emails or stated vocally can be data-mined to determine workers' moods, a process called 'sentiment analysis'. Facial expressions likewise can be analysed to spot signs of fatigue and moods that could be used to make judgements and thus lower OSH risks emerging with overwork. But chatbots, while designed to be assistive machines, still pose psychosocial risks around fears of job loss and replacement. Workers should be trained to understand the role and function of workplace bots and to know what their collaborative and assistive contributions are.

2.4 Wearables and AI in (lot size) manufacturing

Wearable self-tracking devices are increasingly seen in workplaces. The market for wearable devices in industrial and healthcare wearables has been predicted to grow from USD 21 million in 2013 to USD 9.2 billion by 2020 (Nield, 2014). Between 2014 and 2019, 13 million fitness devices are predicted to be incorporated into workplaces. This is already happening in warehouses and factories where GPS, radio-frequency identification and now haptic sensing armbands, such as that patented by Amazon in 2018, have replaced the use of clipboards and pencils.

One new feature of automation and Industry 4.0 processes in which AI-enhanced automation is under way is in the area of lot size manufacturing (10). This process involves workers being provided with glasses with screens and virtual reality functionality, such as HoloLenses and Google Glasses, or computer tablets on stands within the production line, which are used to carry out on-the-spot tasks on production lines. The assembly line model, in which a worker carries out one repeated, specific task for several hours at a time, has not disappeared completely, but the lot size method is different. Used in agile manufacturing strategies, this method involves smaller orders made within specific time parameters, rather than constant bulk production that does not involve guaranteed customers.

In lot size manufacturing, workers are experiencing the introduction of visual on-the-spot training enabled by a HoloLens screen or tablet, in which they are invited to carry out new tasks that are learned instantly and carried out only for the time required to manufacture the specific order a factory receives. While at first glance these assistance systems may *appear* to provide increased autonomy, personal responsibility and self-development, this is not necessarily the case (Butollo, Jürgens and Krzywdzinski, 2018).

The use of on-the-spot training devices, worn or otherwise, means that workers need less pre-existing knowledge or training, because they carry out the work case by case. The risk of work intensification thus arises, as head-mounted displays or tablet computers become akin to live instructors for unskilled workers. Furthermore, workers do not learn long-term skills, because they

(10) Interview with Dr Michael Bretschneider-Hagemes, cited above (fn 2).

are required to perform on-the-spot, modular activities in custom assembly processes, needed to build tailor-made items at various scales. While this is good for the company's efficiency in production, lot size methods have led to significant OSH risks in that that they deskill workers, because skilled labour is needed only to design the on-the-spot training programmes used by those workers who no longer need to specialise themselves.

OSH risks can further emerge because of a lack of communication, meaning that workers are not able to comprehend the complexity of the new technology quickly enough, particularly if they are also not trained to prepare for any hazards arising. One real issue is in the area of small businesses and start-ups, which are quite experimental in the use of new technologies but often overlook ensuring that safety standards are adhered to before accidents occur, when it is, of course, too late ([11]). An interview with those involved in the IG Metall Better Work 2020 project (Bezirksleitung Nordrhein-Westfalen/NRW Projekt Arbeit 2020) revealed that trade unionists are actively speaking to companies about the ways they are introducing *Industry 4.0* technologies into workplaces (Moore, 2018a). The introduction of robots and worker monitoring, cloud computing, machine-to-machine communications and other systems have all prompted those running the IG Metall project to ask companies:

- What impact will technological changes have on people's workloads?
- Is work going to be easier or harder?
- Will work become more or less stressful? Will there be more or less work?

The IG Metall trade unionists indicated that workers' stress levels tended to rise when technologies are implemented without enough training or dialogue with workers. Expertise is often needed to mitigate the risks that new technologies in workplaces introduce.

Next, we turn to another arena in which AI is making an impact, namely in 'gig work' environments.

2.5 Platform applications enabling gig work

'Gig work' is obtained by using online applications (apps), also called platforms, made available by companies such as Uber, Upwork or Amazon Mechanical Turk (AMT). The work can be performed *online* —obtained and carried out on computers in homes, libraries and cafes, for example, and includes translation and design work — or *offline* — obtained online but carried out offline, such as taxi driving or cleaning work. Not all algorithms utilise AI, but the data produced by client-worker matching services and customer assessment of platform workers provide data that train profiles that then result in overall higher or lower scores that then lead, for example, clients to select specific people for work over others.

Monitoring and tracking have been daily experiences for couriers and taxi drivers for many years, but the rise in offline gig workers carrying out platform-directed food delivery by bicycle, delivering orders and taxi services is relatively new. Uber and Deliveroo require workers to install a specific application onto their phones, which perch on vehicle dashboards or handlebars, and they gain clients through the use of mapping satellite technologies and by matching algorithmically operated software. The benefits of using AI in gig work could be driver and passenger protection. DiDi, a Chinese ride-hailing service, uses AI facial recognition software to identify workers as they log on to the application. DiDi uses this information to ensure the identities of drivers, which is seen as a method of crime prevention. However, there was a very serious recent failure in the use of the technology in which a driver logged in as his father one evening. Under the false identity, later in his shift, the driver killed a passenger.

Delivery gig workers are held accountable for their speed, number of deliveries per hour and customer rankings, in an intensified environment that has been proven to create OSH risks. In Harper's magazine a driver explains how new digitalised tools work as a 'mental whip', noting that 'people get intimidated and work faster' (The Week, 2015). Drivers and riders are at risk of

([11]) Prof. Dr Dietmar Reinert, PEROSH Chairman, Institute for Occupational Safety and Health of the German Social Accident Insurance, indicated this in an interview with the author on 13 September 2018.

deactivation from the app if their customer rankings are not high enough or they do not meet other requirements. This results in OSH risks including blatant unfair treatment, stress and even fear.

Algorithms are used to match clients with workers in online gig work (also called micro-work). One platform called BoonTech uses IBM Watson AI Personality Insights to match clients and online gig workers such as those gaining contracts using AMT and Upwork. Issues of discrimination have emerged that are related to women's domestic responsibilities, when carrying out online gig work at home, such as reproductive and caring activities in a traditional context. A recent survey of online gig workers in the developing world conducted by ILO researchers shows that a higher percentage of women than men tend to 'prefer to work at home' (Rani and Furrer, 2017, p. 14). Rani and Furrer's research shows that 32 % of female workers in African countries have small children and 42 % in Latin America. This results in a double burden for women, who 'spend about 25.8 hours working on platforms in a week, 20 hours of which is paid work and 5.8 hours considered unpaid work' (ibid., p. 13). The survey shows that 51 % of women gig workers work during the night (22.00 to 05.00) and the evening (76 % work from 18.00 to 22.00), which are 'unsocial working hours' according to the ILO's risk categories for potential work-related violence and harassment (ILO, 2016, p. 40). Rani and Furrer further state that the global outsourcing of work through platforms has effectively led to the development of a 'twenty-four hour economy … eroding the fixed boundaries between home and work … [which further] puts a double burden on women, since responsibilities at home are unevenly distributed between sexes' (2017, p. 13). Working from home could already be a risky environment for women who may be subject to domestic violence alongside the lack of legal protection provided in office-based work. Indeed, 'violence and harassment can occur … via technology that blurs the lines between workplaces, "domestic" places and public spaces' (ILO, 2017, p. 97).

Digitalising non-standard work, such as home-based online gig work, and taxi and delivery services in offline gig work, is a method of governance that is based on quantification of tasks at a minutely granular level, where only explicit contact time is paid. Digitalisation may appear to formalise a labour market in the ILO sense, but the risk of underemployment and underpay is very real. In terms of working time, preparatory work for reputation improvement and necessary skills development in online gig work is unpaid. Surveillance is normalised but stress still results. D'Cruz and Noronha (2016) present a case study of online gig workers in India, in which 'humans-as-a-service' (as articulated by Jeff Bezos, see Prassl, 2018) is critiqued for being the kind of work that dehumanises and devalues work, facilitates casualisation of workers and even informalises the economy. Online gig work, such as work obtained and delivered using the AMT, relies on non-standard forms of employment (ibid., p. 46) which increases the possibilities of child labour, forced labour and discrimination. There is evidence of racism, whereby clients are reported to direct abusive and offensive comments on the platforms. Inter-worker racist behaviour is also evident: gig workers working in more advanced economies blame Indian counterparts for undercutting prices (ibid.). Further, some of the work obtained on online platforms is highly unpleasant, such as the work carried out by content moderators, who sift through large sets of images and are required to eliminate offensive or distributing images, with very little relief or protection around this. There are clear risks of OSH violations in the areas of heightened psychosocial violence and stress, discrimination, racism, bullying, unfree and underage labour, because of the lack of basic protection in these working environments.

In gig work, workers have been forced to register as self-employed workers, losing out on the basic rights that formal workers enjoy such as guaranteed hours, sick and holiday pay and the right to join a union. Gig workers' online reputations are very important because a good reputation is the way to gain more work. As mentioned above, digitalised customer and client ratings and reviews are key to developing a good reputation and these ratings determine how much work gig workers obtain. Algorithms learn from customer rankings and quantity of tasks accepted, which produces specific types of profiles for workers that are usually publicly available. Customer rankings are deaf and blind to the consideration of workers' physical health, care and domestic work responsibilities, and circumstances outside workers' control that might affect their performance, leading to further OSH risks where people feel forced to accept more work than is healthy, or are at risk of being excluded from work. Customer satisfaction rankings, and number of jobs accepted, can be used to 'deactivate' taxi drivers' use of the platform, as is done by Uber, despite the paradox and fiction that algorithms are absent of 'human bias' (Frey and Osborne, 2013, p. 18).

Overall, there are benefits from integrating AI into gig work, including driver identity protection and allowing flexible hours of work, which are good for people's life and work choices. However, these same benefits can result in rising risks, such as the case of the DiDi driver and the case of a double burden of work for women online workers. OSH protections are generally scarce in these working environments and the risks are many (Huws, 2015; Degryse, 2016) and involve low pay and long hours (Berg, 2016), endemic lack of training (CIPD, 2017) and a high level of insecurity (Taylor, 2017). Williams-Jimenez (2016) warns that labour and OSH laws have not adapted to the emergence of digitalised work, and other studies are beginning to make similar claims (Degryse, 2016). The successes of AI are also its failures.

Having outlined where AI is entering the workplace and the benefits and risks for OSH, the report now turns to look at responses from the wider OSH community, identifying the policy developments, debates and discussions under way on these topics.

3 Policy developments, regulation and training

The emergence of AI and in particular the ecosystem and features of autonomous decision-making require a 'reflection about the suitability of some established rules on safety and civil law questions on liability' (European Commission, 2018). So, horizontal and sectoral rules need to be reviewed to identify any risks arising as well as to protect and ensure benefits from the integration of AI-enhanced technology at work. The Machinery Directive (2006/42/EC), the Radio Equipment Directive (2014/53/EU), the General Product Safety Directive (2001/95/EC) and other specific safety regulations provide some guidance, but more will be needed to ensure workplace safety and health. Indeed, a report in *IOSH Magazine* emphasises that AI risks are 'outpacing our safeguards' (Wustemann, 2017) for workplace safety.

In that light, this chapter looks at the perspectives of policy-makers and experts from the wider community and emerging recommendations for regulating AI to reduce OSH risks, and then it outlines some suggestions for training linked to AI and OSH at IG Metall.

3.1 European Commission

The digital single market is an important vehicle for the expansion of AI, and the mid-term Commission review of the implementation of the digital single market (European Commission, 2017) indicated that AI will provide substantive technological solutions to risky situations such as fewer fatalities on the roads, smarter uses of resources, less use of pesticides, a more competitive manufacturing sector, better precision in surgery and assistance in dangerous situations such as earthquake and nuclear disaster rescue. Surrounding debates across Europe involve questions about the legal and liability issues, data sharing and storage, the risks of bias in machine learning's competences and the difficulty of allowing for the right to an explanation, including how data about workers are used, firmed up by the General Data Protection Regulation (GDPR).

So, the arena that is covered by the mid-term review of the digital single market, which has implications for AI, OSH and work, is the discussion of the risks of bias and the right to an explanation of how data are used, in which informed consent for the use of data and the right to access data held about the person is paramount. The socio-economic and ethical issues of AI have been further highlighted in more recent European Commission Communications, particularly since the April 2018 Communication on artificial intelligence for Europe and conclusions on the coordinated plan for the development and use of artificial intelligence made in Europe, which emphasise ethics for competitive advantage.

3.2 International standards

A committee within the International Organization for Standardization (ISO) in 2018 and 2019 has been working on designing a standard that will apply to uses of dashboards and metrics in workplaces. The standard will include regulations on how dashboards can be set up and on gathering and using data from workers. Quantification tools are becoming increasingly of interest to employers, but the data are of no use if they are not standardisable. Representatives from the manufacturers of the software used to standardise data, SAP, are active in the ISO discussions, but it will be important for further actors to participate — in Germany, for example, IG Metall is looking at training and AI — in order to ensure co-determination and to ensure workers' representation more broadly across the international landscape. One expert in this area indicated that international standards can be an effective way to ensure that the benefits of these tools are achieved, and an important step is to ensure that international corporate practices are equivalent at some level, that data are standardisable and that workers are involved in the discussions and implementation processes ([12]). Furthermore, risks assessments should be carried out on the basis of the extensive data gathered by these dashboards, which are a clear benefit for OSH protection.

3.3 International Labour Organization

The ILO has produced a number of reports suggesting best practice in integrating AI into workplaces, for member states.

In the report named Digital Labour Platforms and the Future of Work (Berg et al., 2018), the human side of AI is revealed, whereby many of the jobs or micro-tasks obtained with online platform work (outlined above) are similar to unskilled work that can also be, in many cases, automated. The report indicates that micro-task platforms were in fact invented in part to deal with the failures of algorithms of Web 2.0 to 'classify the nuances of the images, sounds and texts' that companies wanted to store and classify (Irani, 2015, p. 225, cited in Berg et al., 2018, p. 7). Work can range from bulk tasks, such as a survey requiring thousand of replies, to image recognition. Amazon actually calls the kind of work performed by people using its AMT platform 'artificial artificial intelligence', or 'an on-demand, scalable, human workforce to complete jobs that humans can do better than computers, for example, recognising objects in photos' (Berg et al., 2018, p. 7).

The report recommends regulation of crowdwork platforms (which it argues replace AI and automated work, as indicated above), listing 18 criteria for 'fairer microwork', which include such recommendations as eliminating misclassification as 'self-employed' when workers are employees in practice; rights to union membership and collective bargaining; minimum wage; fee transparency (wage theft is a common problem in gig work); allowing workers to accept some tasks and decline others and not be penalised for it; protections from computer failure; readable and concise platform terms; protections from misuse of worker evaluations and ratings; available codes of conduct; permitting workers to contest non-payment and other issues; giving workers access to information about clients; platforms reviewing task instructions before posting; enabling workers to export reputation histories; giving workers the right to work with a client after working via the platform; customers and operators responding to worker requests promptly and politely; workers knowing the purpose of their work; and any tasks involving psychologically stressful work being labelled clearly (Berg et. al., 2018, pp. 105-109).

The Global Commission on the Future of Work 'Work for a brighter future' report indicates that any actions involving technology and work must take a human-centred agenda. Cobots, the report notes, can actually reduce worker stress and risks of injury. However, technology can also reduce the availability of work for humans, which ultimately will alienate workers and stunt workers' development. Workplace decisions should never rely on data produced by algorithms, and any AI at work should be integrated with a 'human-in-command' approach, whereby any 'algorithmic management, surveillance and control, through sensors, wearables and other forms of monitoring,

([12]) Interview with Rolf Jaeger, European Industrial Relations Intercultural Communication and Negotiation, 18 September 2018.

needs to be regulated to protect the dignity of workers' (ILO, 2019, p. 43). The report goes on to state, building on the ILO Declaration of Philadelphia statement that labour is not a commodity: 'Labour is not a commodity; nor is it a robot' (ibid.).

3.4 World Economic Forum and GDPR

The World Economic Forum (WEF) Global Future Council on Human Rights and Technology reported in 2018 that, even when good data sets are used to set up algorithms for machine learning, there are considerable risks of discrimination if the following occur (WEF, 2018):

1. choosing the wrong model;
2. building a model with inadvertently discriminatory features;
3. absence of human oversight and involvement;
4. unpredictable and inscrutable systems;
5. unchecked and intentional discrimination.

The WEF emphasises that there is a distinct need for 'more active self-governance by private companies', which is in line with the ILO's Tripartite Declaration of Principles concerning Multinational Enterprises and Social Policy — 5th Edition (Rev. 2017), which provides direct guidance for enterprises in the areas of sustainable, responsible and inclusive working practices and social policy surrounding these, whereby the Sustainable Development Goal (SDG) target 8.8 aims to achieve safe and secure working environments for all workers by 2030. The prevention of unfair and illegal discrimination clearly must be ensured as AI is increasingly introduced, and the WEF (2018) and ILO reports mentioned above are vital for steering the course.

The first error a company can make when using AI, which could lead to discrimination as listed by the WEF, is in a situation in which a user applies the same algorithm to two problems that may themselves not have identical contexts or data points. A possible workplace example of this could be where potential hires are considered using an algorithm that looks for clues about personality types via searches through social media, videos that detect facial movements and data that are collated across data sets of curriculum vitae, perhaps extending back a few years of hiring. As Dr Cathy O'Neil pointed out in an interview with the author (13), the algorithm then has to be designed to be discriminatory or at least selective, because hiring practices require that at a basic level. However, if, for example, the algorithm is looking for extroverted people for a call centre job, the same algorithm would not be appropriate to find the right lab assistant, where talkativeness is not inherent to the job description. While applying the algorithm would not necessarily lead to illegal discrimination as such, it is not difficult to extrapolate the possibilities for misallocation.

The second error, 'building a model with inadvertently discriminatory features', can refer to, for example, using a databank that already exemplifies discrimination. For example, in the United Kingdom, the gender pay gap has been exposed recently, revealing that for years women have been working for lower salaries and in some cases doing the same work as men for less pay. If the data that demonstrate this trend were used to create an algorithm to make decisions about hiring, the machine would 'learn' that women should be paid less. This demonstrates the point that machines cannot make ethical judgements independently of human intervention. Indeed, there is a growing arena of research that demonstrates that discrimination is not eliminated by AI in decision-making and prediction but instead that codification of data perpetuates the problem (Noble, 2018).

The third error emphasises human intervention, which is now required across Europe. In May 2018, the GDPR became a requirement, whereby workers' consent for data collection and use applies. While the GDPR looks primarily at consumer data rights, there are significant applications for the workplace, because workplace decisions cannot be made using automated processes alone.

([13]) Dr Cathy O'Neil, author of *Weapons of maths destruction* and CEO of ORCAA (O'Neil Risk Consulting and Algorithmic Auditing), interviewed by the author on 14 October 2018.

Section 4 of the GDPR outlines the 'Right to object and automated individual decision-making'. Article 22, 'Automated individual decision-making, including profiling', indicates that:

> 22(1): The data subject shall have the **right not to be subject to a decision based solely on automated processing**, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.

The foundations for the Regulation, listed in the first sections of the document, make it clear that:

> (71): The data subject has the **right not to be subject to a decision**, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing and which produces legal effects concerning him or her or similarly significant affects him or her, such as… **e-recruiting practices without any human intervention**. Such processing includes profiling that consists of any form of automated processing of personal data evaluating the personal aspects of a natural person, in particular to analyse or predict aspects concerning the data subject's performance at work… reliability or behaviour, location or movements, where it produces legal effects concerning him or her or similarly significantly affects him or her.

Failure to apply these criteria could lead to unfair or illegal discriminatory decisions.

With regard to the fourth error, 'unpredictable and inscrutable systems', the description in the WEF (2018) report indicates that 'when a human makes a decision, such as whether or not to hire someone, we can inquire as to why he or she decided one way or the other'. Obviously, a machine cannot discuss its 'rationale' for decisions it reaches based on data mining. The elimination of qualified judgements and lack of human intervention therefore creates a clear route to discrimination.

The final error in the implementation of AI can be when 'unchecked and intentional discrimination' occurs. This could happen, for example, when a company in fact does not want to hire women who are likely to become pregnant. While this explicit position would not hold up in court, a machine learning system could provide covert tactics to make it happen through an algorithm designed to out to filter out a subset of women candidates where that might be the case, based on age and relationship data. It is not difficult to see how this opens the door to not only the risks, but in fact the likelihood, of technically illegal discrimination.

3.4 Training for AI and OSH

IG Metall is working with companies on its OSH training programmes to accommodate the latest technological changes in workplaces in 2019. Discussions with the expert leading this initiative indicated that training for OSH has typically been seen as an arena solely populated by the one or two safety and health officers in workplaces and has not been fully integrated into all systems. Findings now demonstrate that people need to be trained to acquire fast learning capabilities, because technology changes quickly and skills thus must adapt (14). This expert indicated that training must be adjusted for relevance in the era of Industry 4.0 and digitalisation, so that workers are prepared to deal with emerging risks. However, this is not a panacea and must be part of a larger implementation plan. If there is no plan to actually implement and use any new knowledge and skills delivered by training, new skills will be lost. In that light, better alignment between OSH training and integrated technologies is necessary. That being said, training pedagogy should also be adjusted, as learning is a process that will need to continue throughout workers' lifetimes, particularly in the current climate of job uncertainty. It will also be important for workers to acquire problem-solving skills and principles as well as traditionally understood 'skills'. These days,

(14) Interview with Dr Maike Pricelius, Project Secretary, Better Work 2020, IG Metall, 12 October 2018.

workers should understand and choose their own learning pathways and styles ([15]). Only time will tell just how ubiquitous AI in workplaces becomes, but it is worth remaining alert to OSH risks and benefits and involving workers in these processes by providing training at every juncture.

4 In conclusion

Even as far back such as the 1920s, the writer E. M. Forster painted a dystopian picture of technology and humanity. Forster's classic story *The machine stops* talks about a world where humans must live underneath the Earth's surface, within a machinic apparatus that the protagonist of this novel celebrates, because the machine (Forster, 1928):

> … feeds us and clothes us and houses us; through it we speak to one another, through it we see one another, in it we have our being. The Machine is the friend of ideas and the enemy of superstition: the Machine is omnipotent, eternal; blessed is the Machine!

But the omnipotent, all-housing contraption soon begins to decay in this classic literary masterpiece, and human expertise is not sufficient for its maintenance, leading to a grim ending for all of humanity.

While this is a classic piece of science fiction, today, technology's seeming invisibility and potential power are seemingly endlessly perpetuated, because its operations are often hidden within a black box, where its workings are often considered beyond comprehension but seem, still, to be accepted by the majority of people. Most people are not engineers, so they do not understand how computers and AI systems work. Nonetheless, human experts are surprised by AI actions, such as the chess or Go player who was beaten by a computer programme.

In China, the government will soon give each person a citizen score, or an economic and personal reputation scoring, which will look at people's rent payments, credit rankings, phone use, and so on. It will be used to determine conditions for obtaining loans, jobs and travel visas. Perhaps people analytics could be used to give people 'worker scores' to be used for decision-making in appraisals, which would introduce all sorts of questions about privacy and surveillance. The 'algorithmic condition' is a term also coined in a recent EU report (Colman et al., 2018) that refers to the increasingly normalised logic of algorithms, in which symbols are transformed into reality. Today, this condition is beginning to impact many workplaces, in which online reputations are subject to algorithmic matching and people's profiles subject to data-mining bots. The problem is that algorithms do not see the qualitative aspects of life or the surrounding contexts. Dr O'Neil (cited in fn 13) made an insightful observation in a recent interview with the author. While watching Deliveroo riders hurtle past her in the rain, Dr O'Neil considered the platforms directing the riders' work, which operate on the basis of efficiency and speed and thus instigate riders to cycle fast in unsafe weather conditions. This clearly puts riders' very lives at risk. Dr O'Neil calls algorithms 'toy models of the universe', because these seemingly all-knowing entities actually only know what we tell them, and thus they have major blind spots.

Google co-founder Sergey Brin addressed investors in his annual founder's letter earlier in 2018 stating that:

> …the new spring in AI is the most significant development in computing in my lifetime … however, such powerful tools also bring with them new questions and responsibilities. How will they affect employment across different sectors? How can we understand what they are doing under the hood? What about measures of fairness? How might we manipulate them? Are they safe?

Ethical questions in AI must be discussed beyond the corporate sphere, however, and this report has covered these questions surrounding OSH and the risks as well as benefits that are introduced. The mythical invention of E. M. Forster's all-encompassing machine in his classical science fiction story was not, of course, subject to a range of ethical and moral review panels before all of

([15]) Duncan Spencer, Head of Advice and Practice, Institution of Occupational Safety and Health, based in Leicester, United Kingdom, talked about this with the author in an interview on 15 October 2018.

humanity began to live within it under the Earth's crust. This dystopia is not exactly what we face now, of course, but current discussions — from those held to inform the European Commission Communications and European Coordinated Plan on Artificial Intelligence to trade union curriculum review groups such as those at IG Metall — show significant interest in preventing the worst risks and facilitating the benefits of OSH, as AI is increasingly implemented for workplace decision-making and assisted work.

In conclusion, as the implementation of AI at work is relatively new, there is only nascent evidence of OSH risks and benefits. Nonetheless, this report has covered some of the arenas where benefits are being noted and supported and risks are highlighted, as well as caution and regulation being applied. In HR decision-making with the use of AI-augmented people analytics, the risk of unfair treatment and discrimination is flagged up. In automation and Industry 4.0, risks involve unsuitable or unavailable training leading to overwork and stress (Downey, 2018), unpredicted accidents such as collisions between humans and robots. Deskilling of work is at stake in manufacturing and other industries, with the integration of lot size processes and the use of wearable technologies for automated training practices. Risks of privacy relating to intensified surveillance and feelings of micro-management have been reported, as management is able to access more intimate data about workers because of wearable technologies in factories and offices alike. In gig work, algorithms cannot be posed as sole decision-makers. The benefits should be emphasised in all cases.

Indeed, it will be important for all stakeholders to maintain focus on the assistive possibilities of business application and ensure government and other regulatory oversight of AI tools and applications in the workplace. The positive effects of AI, when it is implemented with appropriate processes, are that it can help management reduce human bias in interviewing processes, if algorithms are designed to identify evidence of past discrimination in decision-making and decisions are made with full human intervention and even affirmative action. AI can help to improve relationships with, and between, employees when data collected demonstrates potential for collaboration. AI-enhanced HR tools can improve decision-making using prediction by exception and can allow people more time for personal and career development, if AI can start to take over repetitive and unfulfilling work.

To avoid risks for OSH, the author recommends focusing on implementing assistive and collaborative AI rather than heading for the general and widespread competences of universal AI. Appropriate training must be provided at all points, and checks must also be made consistently, including by OSH departments and authorities. Workers must be consulted at all points whenever new technologies are integrated into workplaces, sustaining a worker-centred approach and prioritising a 'human in command' approach (De Stefano, 2018). Business owners and governments should keep an eye on international standardisation, government regulation and trade union activities, in which significant progress is already being made to mitigate against the worst risks of AI and to develop positive and beneficial gains. In conclusion, it is not AI technology itself that creates risks for the safety and health of workers, it is the way that it is implemented, and it is up to all of us to ensure a smooth transition to increased integration of AI in workplaces.

Author: Dr Phoebe V. Moore, Associate Professor of Political Economy and Technology, Management and Organisation Division, School of Business, University of Leicester, United Kingdom, and Research Fellow, WZB Weizenbaum Institute for the Networked Society 2018-19

Project management: Annick Starren — European Agency for Safety and Health at Work (EU-OSHA)

References

Agarwal, A., Gans, J., Goldfarb, A., 2018, *Prediction machines: The simple economics of artificial intelligence*, Boston, MA: Harvard Business Review Press.

Berg, J., 2016, *Income security in the on-demand economy: Findings and policy lessons from a survey of crowdworkers*, Conditions of Work and Employment Series No 74, Geneva: International Labour Organization.

Berg, J., Furrer, M., Harmon, E., Rani, U., Silberman, M. S., 2018, *Digital labour platforms and the future of work: Towards decent work in the online world*, Geneva: International Labour Organization.

Butollo, F., Jürgens, U., Krzywdzinski, M., 2018, 'From lean production to Industrie 4.0: More autonomy for employees?', Wissenshanftszentrum Berlin für Socialforschung (WZB) Discussion Paper SP 111 2018-303.

CIPD (Chartered Institute for Personnel Development), 2017, *To gig or not to gig? Stories from the modern economy*. Available online: www.cipd.co.uk/knowledge/work/trends/gig-economy-report

Collins, L., Fineman, D. R., Tshuchica, A., 2017, 'People analytics: Recalculating the route', Deloitte Insights. Available online: https://www2.deloitte.com/insights/us/en/focus/human-capital-trends/2017/people-analytics-in-hr.html

Colman, F., Bülmann, V., O'Donnell, A., van der Tuin, I., 2018, *Ethics of coding: A report on the algorithmic condition*, Brussels: European Commission.

D'Cruz, P., Noronha, E., 2016, 'Positives outweighing negatives: The experiences of Indian crowdsourced workers', *Work Organisation, Labour and Globalisation* 10(1), 44-63.

De Stefano, V., 2018, 'Negotiating the algorithm: Automation, artificial intelligence and labour protection', ILO working Paper No 246/2018, Geneva: International Labour Organization.

Degryse, C., 2016, *Digitalisation of the economy and its impact on labour markets*, Brussels: European Trade Union Institute (ETUI).

Delponte, L., 2018, *European artificial intelligence leadership, the path for an integrated vision*, Brussels: Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament.

Downey, K., 2018, 'Automation could increase workplace stress, unions warn', *IOSH Magazine*, 23 April 2018. Available online: https://www.ioshmagazine.com/article/automation-could-increase-workplace-stress-unions-warn

Dreyfus, H. L., 1972, *What computers can't do*, New York: Harper and Row (reprinted by MIT Press 1979, 1992).

Ernst, E., Merola, R., Samaan, D., 2018, *The economics of artificial intelligence: Implications for the future of work*, ILO Future of Work Research Paper Series, Geneva: International Labour Organization. Available online: https://www.ilo.org/global/topics/future-of-work/publications/research-papers/WCMS_647306/lang--en/index.htm

EU-OSHA (European Agency for Safety and Health at Work), 2018, *Foresight on new and emerging occupational safety and health risks associated with digitalisation by 2025*, Luxembourg: Publications Office of the European Union. Available online: https://osha.europa.eu/en/tools-and-publications/publications/foresight-new-and-emerging-occupational-safety-and-health-risks/view.

European Commission, 2018, Communication on artificial intelligence for Europe, Brussels: European Commission.

European Commission, 2017, Communication on the mid-term review on the implementation of the digital single market strategy: A connected digital single market for all. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1496330315823&uri=CELEX:52017DC0228

Feloni, R., 2017, 'I tried the software that uses AI to scan job applicants for companies like Goldman Sachs and Unilever before meeting them, and it's not as creepy as it sounds', Business Insider UK, 23 August 2017. Available online: https://www.uk.businessinsider.com/hirevue-ai-powered-job-interview-platform-2017-8?r=US&IR=T/#in-recorded-videos-hirevue-employees-asked-questions-like-how-would-you-describe-your-role-in-the-last-team-you-worked-in-4

Forster, E. M., 1928/2011, *The machine stops*, London: Penguin Books.

Frey, C., Osborne, M. A., 2013, *The future of employment: How susceptible are jobs to computerisation?*, Oxford: University of Oxford, Oxford Martin School. Available online: https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

Houghton, E., Green, M., 2018. *People analytics: Driving business performance with people data*, Chartered Institute for Personnel Development (CIPD). Available online: https://www.cipd.co.uk/knowledge/strategy/analytics/people-data-driving-performance

Hutter, M., 2012, 'One decade of universal artificial intelligence', *Theoretical Foundations of Artificial General Intelligence* 4, 67-88.

Huws, U., 2015, 'A review on the future of work: Online labour exchanges, or "Crowdsourcing" — Implications for occupational safety and health', Discussion paper, Bilbao: European Agency for Safety and Health at Work. Available online: https://osha.europa.eu/en/tools-and-publications/publications/future-work-crowdsourcing/view

IBM, 2018, 'IBM talent business uses AI to rethink the modern workforce', IBM Newsroom. Available online: https://newsroom.ibm.com/2018-11-28-IBM-Talent-Business-Uses-AI-To-Rethink-The-Modern-Workforce

ILO (International Labour Organization), 2019, *Work for a brighter future: Global Commission on the Future of Work*, Geneva: International Labour Organization.

ILO (International Labour Organization), 2017, 'Ending violence and harassment against women and men in the world of work, Report V', International Labour Conference 107th Session, 2018, Geneva. Available online: http://www.ilo.org/ilc/ILCSessions/107/reports/reports-to-the-conference/WCMS_553577/lang--en/index.htm

ILO (International Labour Organization), 2016, *Final Report: Meeting of Experts on violence against women and men in the world of work*, MEVWM/2016/7, Geneva: ILO. Available online: http://www.ilo.org/gender/Informationresources/Publications/WCMS_546303/lang--en/index.htm

Kaivo-oja, J., 2015, 'A review on the future of work: Robotics', Discussion paper, Bilbao: European Agency for Safety and Health at Work. Available online: https://osha.europa.eu/en/tools-and-publications/seminars/focal-points-seminar-review-articles-future-work

Kar, S., 2018, 'How AI is transforming HR: The future of people analytics', Hyphen, 4 January 2018. Available online: https://blog.gethyphen.com/blog/how-ai-is-transforming-hr-the-future-of-people-analytics

McCarthy, J., Minsky, M. L., Rochester, N., Shannon, C. E., 1955, 'A proposal for the Dartmouth Summer Research Project on Artificial Intelligence'. Available online: http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

Moore, P. V., 2018a, *The threat of physical and psychosocial violence and harassment in digitalized work*, Geneva: International Labour Organization.

Moore, P. V., 2018b, *The quantified self in precarity: Work, technology and what counts*, Abingdon, UK: Routledge.

Nield, D., 2014, 'In corporate wellness programs, wearables take a step forward', *Fortune*, 15 April 2014. Available online: http://fortune.com/2014/04/15/in-corporate-wellness-programs-wearables-take-a-step-forward/

Noble, S. A., 2018, *Algorithms of oppression: How search engines reinforce racism*, New York: New York University Press.

Pasquale, F., 2015, *The black box society: The secret algorithms that control money and information*, Boston, MA: Harvard University Press.

Prassl, J., 2018, *Humans as a service: The promise and perils of work in the gig economy*, Oxford: Oxford University Press.

PwC, 2018a, 'Artificial intelligence in HR: A no-brainer'. Available online: https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html

PwC, 2018b, 'AI will create as many jobs as it displaces by boosting economic growth'. Available online: https://www.pwc.co.uk/press-room/press-releases/AI-will-create-as-many-jobs-as-it-displaces-by-boosting-economic-growth.html

Rani, U., Furrer, M., 2017, 'Work and income security among workers in on-demand digital economy: Issues and challenges in developing economies', Paper presented at the Lausanne University workshop 'Digitalization and the Reconfiguration of Labour Governance in the Global Economy', 24-25 November 2017 (unpublished).

Simon, H., 1969, *The sciences of the artificial*, Cambridge, MA: MIT Press.

Taylor, M., 2017, *Good work: The Taylor review of modern working practices*, London: Department for Business, Energy and Industrial Strategy. Available online: https://www.gov.uk/government/publications/good-work-the-taylor-review-of-modernworking-practices

The Week, 2015, 'The rise of workplace spying', Available online: http://theweek.com/articles/564263/rise-workplace-spying

TNO (Netherlands Organisation for Applied Scientific Research), 2018, *Emergent risks to workplace safety; Working in the same space as a cobot*, Report for the Ministry of Social Affairs and Employment, The Hague.

Turing, A. M., 1950, 'Computing machinery and intelligence', *Mind* 49, 433-460.

United Kingdom (UK) Department for Business, Energy and Industrial Strategy and Department for Digital, Culture, Media and Sport, 2018, 'AI sector deal policy paper'. Available online: https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal

WEF (World Economic Forum), 2018, *How to prevent discriminatory outcomes in machine learning*, World Economic Forum Global Future Council on Human Rights 2016-2018, Cologny, Switzerland: WEF.

Weizenbaum, J., 1976, *Computer power and human reason: From judgment to calculation*, San Francisco: W. H. Freeman.

White House Office of Science and Technology Policy, 2018, *Summit on artificial intelligence for American industry*. Available online: https://www.whitehouse.gov/articles/white-house-hosts-summit-artificial-intelligence-american-industry/ Summary of report. Online: https://www.whitehouse.gov/wp.../Summary-Report-of-White-House-AI-Summit.pdf

Williams-Jimenez, I., 2016, 'Digitalisation and its impact on psychosocial risks regulation', Paper presented at the Fifth International Conference on Precarious Work and Vulnerable Workers, London, Middlesex University.

Woodcock, J., 2016, *Working the phones: Control and resistance in call centres*, London: Pluto Press.

Wustemann, L., 2017, 'AI and nanotech risk outpacing our safeguards', *IOSH Magazine*, 25 August 2017. Available online: https://www.ioshmagazine.com/article/ai-and-nanotech-risk-outpacing-our-safeguards