

My boss the algorithm: an ethical look at algorithms in the workplace

6 March 2020

Contents

[Disclaimer](#)

[Executive summary](#)

[Introduction](#)

[Areas of use](#)

[Opportunities and risks](#)

[Towards more ethical algorithms](#)

[Bibliography](#)

Disclaimer

This is an independent, evidence-based policy paper prepared for Acas by Patrick Briône from the [Involvement and Participation Association](#) (IPA). The views in this paper are the author's own and do not necessarily reflect those of Acas nor the Acas Council.

This paper is not intended as guidance from Acas about using algorithmic management. It's also not intended as an endorsement by Acas of practices to be adopted in the workplace.

Executive summary

Algorithms are becoming more widespread in many parts of our lives. The workplace is no exception, with a rise in what has been termed 'algorithmic management'. This presents new opportunities to improve workplace outcomes, as well as new concerns and risks for workers. This report seeks to examine this trend from both a practical and ethical perspective, in an attempt to provide answers to the following important questions:

Where are algorithms being used in the workplace and how widespread is their adoption?
How can algorithms be used to improve workplace outcomes?

What are the unintended consequences that algorithms can have and how can they worsen workplace outcomes?

What future risks and opportunities of the growing use of algorithmic management can be identified? Do they have the potential to significantly change the employment relationship?

How should responsible employers approach the use of algorithms in the workplace and discussion about them with the workforce?

The introduction to this report sets out the definitions of algorithms, artificial intelligence (AI) and machine learning and the different kinds of algorithmic management taking place in the UK. Specifically, this report has identified and explored 3 main areas of use: algorithmic recruitment, algorithmic task-allocation and algorithmic monitoring and performance review of the workforce. These are each explored in detail in their own chapters.

The report then turns to look at the opportunities and risks that algorithmic management presents. There are 2 obvious benefits on offer – improved productivity through time saved and more efficient decision-making; and new insights into workplace behaviour, human relationships or other trends as a result of vast data processing that can enable whole new solutions to workplace problems.

There are 2 obvious risks and drawbacks – firstly, a threat of increased management control without corresponding consent from the workforce, particularly in areas of surveillance and performance monitoring. Also, there is a danger of eroding human autonomy by replacing the personal relationships of line managers and their reports with a dehumanized system of being managed by a machine.

Finally, there are 2 related areas that could be both opportunities and risks – the impact of algorithms on increasing or reducing bias, and on increasing or reducing accuracy of decision-making. Depending on how the technology is used and how suitable the tasks are to which it is allocated, algorithmic management has the potential to greatly improve or greatly worsen outcomes on either of those fronts.

This report concludes with a look at the ethics of algorithmic management and the approach responsible employers should take when considering these tools, if they want to maximize the opportunities and minimise the risks. Our recommendations are summarised below.

Recommendations

Algorithms should be used to advise and work alongside human line managers but not to replace them. A human manager should always have final responsibility for any workplace decisions.

Employers should understand clearly the problem they are trying to solve and consider alternative options before adopting an algorithmic management approach.

Line managers need to be trained in how to understand algorithms and how to use an ever-increasing amount of data.

Algorithms should never be used to mask intentional discrimination by managers.

There needs to be greater transparency for employees (and prospective employees) about when algorithms are being used and how they can be challenged; particularly in recruitment, task allocation and performance management.

We need agreed standards on the ethical use of algorithms around bias, fairness, surveillance and accuracy.

Early communication and consultation between employers and employees are the best way to ensure new technology is well implemented and improves workplace outcomes.

Company reporting on equality and diversity, such as around the gender pay gap, should include information on any use of relevant algorithms in recruitment or pay decisions and how they are programmed to minimise biases.

The benefits of algorithms at work should be shared with workers as well as employers.

We need a wider debate about the likely winners and losers in the use of all forms of technology at work.

Introduction

Understanding algorithms

An algorithm, in its simplest definition, is just a list of logical instructions that set out how to accomplish a particular task – a way of mapping certain inputs to particular outputs. In a sense anything from a set of assembly instructions for IKEA furniture to a cake recipe could be considered an algorithm. However, the term is usually used to refer to digitalized algorithms embedded in computer code.

"They take a sequence of mathematical operations – using equations, arithmetic, algebra, calculus, logic and probability – and translate them into computer code. They are fed with data from the real world, given an objective and set to work crunching through the calculations to achieve their aim." (Fry, 2018)

Algorithms, AI and machine learning

While some algorithms are quite simple, such as an algorithm that tracks what time an employee first logs on to their work computer each day, or that automatically emails shift workers to remind them of their schedule the day before they are next due in, others can be much more complex. Artificial Intelligence, or AI, is a term that usually refers to a combination of many algorithms working together that have the ability to edit themselves to try and improve their own function.

Some of the most advanced use what is known as machine learning or neural nets, allowing them to learn for themselves how to achieve a set goal rather than just following pre-programmed steps. At the extreme end machine learning systems, like IBM's Watson or GPT-2, can be set to a very wide range of tasks: write essays, news articles or poetry, detect cancers, win quiz shows and even play chess without being taught how to play.

For simplicity in this report we use the word 'algorithms' to describe all these different kinds of systems, but it is worth remembering the variety of technologies and range of capabilities that this term can cover.

The phrase 'algorithmic management' was first coined by a group of academics describing the ways gig economy platforms such as Uber and Lyft used software algorithms to allow workers to be "assigned, optimized and evaluated through algorithms and tracked data" (Lee, Kusbit, Metsky and Dabbish, 2015).

According to a special Data & Society report on the subject, "Algorithmic management is a diverse set of technological tools and techniques to remotely manage workforces, relying on data collection and surveillance of workers to enable automated or semi-automated decision-making." (Mateescu and Nguyen, 2019)

While many of the technologies of algorithmic management originated in the gig economy, "algorithmic management is becoming more common in other work contexts beyond 'gig' platforms" and is being adopted in many other workplace contexts.

It is typically defined to include the following aspects:

- prolific data collection and surveillance of workers through technology
- real-time responsiveness to data that informs management decisions
- automated or semi-automated decision-making
- transfer of performance evaluations to rating systems or other metrics
- the use of 'nudges' and penalties to indirectly incentivize worker behaviours (Mateescu and Nguyen, 2019)

Other terms are also used to describe similar sets of technologies. Digitalized management methods (DMMs) is a catch-all term that can be used to describe the use of algorithms, AI and other digital tools in distributing and reducing work; reorganizing and tracking locations of work; recruiting, appraising and firing workers; accelerating standards and targets; and monitoring and tracking productivity. (Moore and Joyce, Black box or hidden abode? Control and resistance in digitalized management, 2018). This can cover both the use of such technology in regular workplaces as well as its use as the foundations of the platform-based gig economy.

Another related term is that of 'people analytics', which can be broadly defined as the use of statistical tools such as algorithms, big data and AI to 'measure, report and understand employee performance, aspects of workforce planning, talent management and operational management' (Collins, Fineman, and Tsuchida, 2017). This research also shows that 71% of international companies consider people analytics to be a high priority for their organisations.

While there are clear potential benefits to firms employing these kinds of technologies in terms of improved efficiency, speed of decision making and ability to coordinate the behaviour of large numbers of workers and customers, there are clear potential risks for employees if the technology is misused.

"There is growing evidence, however, that digitalized management methods (DMMs) themselves put people into situations where the risks of PPVH [physical and psychosocial violence and harassment] are high." (Moore, 2018)

Other major areas of concern are around increases in surveillance and control, a potential lack of transparency, the risks of hiding and embedding bias and discrimination and a lack of accountability when decision-making becomes detached from human managers.

Prevalence and adoption

According to an EU report, "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." (European Agency for Safety and Health at Work, 2019)

Of the top 100 largest AI start-ups in the world, 3 – US firms Mya Systems and Textio and Israel-based Workey – are focused on HR tech (Rapp and O'Keefe, 2018). Many other smaller AI and tech companies are moving into the HR space as well.

According to the latest figures, "one estimate suggests a fifth of employers in Europe had access to wearable tech in 2015, while in the US as many as 72% of CVs are not seen by human eyes. Amazon, Unilever, Deloitte, Tesco – nearly every major corporate has dipped their toe in the water of algorithmic management." (Dellot, 2017)?

In the UK the NHS has partnered with Humanyze to trial their wearable data-gathering technology which analyses how workers communicate and relate to each other, their physical movements around the workplace and their psychological state.

The most common usages of algorithmic management can be found in 3 areas of business practice. Firstly, recruitment, where it includes at the most basic level CV screening, but can extend to psychometric testing, automated interviewing and even facial recognition and expression analysis.

Secondly, in the day-to-day management functions of task allocation or shift allocation; making decisions about what workers should be doing and when. Automatic shift allocation software is becoming extremely prevalent in the retail and hospitality sectors, while manufacturing and logistics firms are using algorithms to micro-manage in ever greater detail the individual movements and actions of their workers on a minute-by-minute basis.

Finally, there is the growth of performance review algorithms; those which are designed not to give instructions to workers but to collect data on them and feed it back to managers, who can use the outputs to make decisions that could include pay, promotion or firing.

Investment in such technologies has grown apace and is being driven in part by what academics describe as an 'arms race' – that companies will purchase any technology that looks shiny and offers the prospect of getting ahead of, or, at least keeping pace with their competitors, even if they don't understand fully what it does or whether it suits their needs.

This arms race is being led by those at the top of the tech industry such as IBM, or Google who in their 2017 website providing guidance on people analytics states that a data-driven approach to HR management is the best way to "inform your people practices, programs and processes... predictive analytics uncover new insights, solve people problems and direct HR actions." (Google, 2017)

As Dr Phoebe Moore points out in the International Labour Organization (ILO) report into the subject, "'People problems' could, of course, mean 'who to fire' or decisions on 'who not to promote' and the like. In any case, without human intervention, these HR judgements become potentially very dubious when the qualitative dimensions of the workplace are eliminated, and could increase workers' stress." (Moore, 2018)

Benefits and drawbacks

The most important point to grasp is that algorithms are tools. Like other tools from simple hammers to vast and complex systems like the internet, algorithms are neither inherently good nor inherently bad. Rather they multiply the potential of human beings to achieve both good and bad outcomes.

Just as the use of algorithms in healthcare or policing are offering transformative benefits in those fields, so algorithmic management could have a transformative effect on many areas of the workplace. Major potential benefits include improved accuracy of decision making, reduced time and cost for both managers and workers, less bias leading to more impartial decision making and new data-driven insights into their workforce that allow organisations to do things that were not previously possible.

Examples of all these potential benefits are described in more detail throughout this paper:

- more efficient shift scheduling meaning less wasted time for both managers and workers
- more efficient task allocation in factories meaning increased productivity
- better performance assessments through more accurate data collection
- improved speed of recruitment and better candidate quality
- reduced opportunities for human favouritism and unconscious biases to intrude into management decisions around remuneration, holiday approval or shift allocation decisions

Of course, all of these are potential benefits. In practice algorithms can often be flawed in various ways, such that they sometimes have the opposite effects to those intended; amplifying rather than eliminating biases and reducing rather than increasing accuracy,

not to mention certain additional drawbacks of risking dehumanising the management process and alienating workers. The opportunities for additional surveillance and control of the workforce that these tools provide could also prove to be dangerous temptations for managers to overstep ethical boundaries.

The same performance monitoring tool that identifies poor performers could be used by a good manager to provide extra targeted support to workers that need it, or by a bad manager to simply get rid of the bottom 10% of their workforce. The algorithm in either case is the same, but it is the use to which it is put by human managers that makes the ethical difference.

Similarly, where algorithms end up embedding and perpetuating biases rather than removing them, this isn't generally an inherent feature of using AI but, as discussed below, more to do with the variables it's told to use in its decision making and the training data it is fed being full of past examples of biased human decision making. In other words, the fault, often, is not in our algorithms, but in ourselves.

Areas of use

Use of algorithms in recruitment

There are several distinct stages of the recruitment process for any new candidate in a typical job. Firstly, the choice of where and how to advertise the posting, followed by a screening of CVs to create a shortlist for interview; then the interview process itself, possibly accompanied by additional testing or assessments; and then finally selecting and onboarding the successful candidate. Increasingly, algorithms are coming to play a significant role in every single step of this process.

While many jobs are still advertised through general postings on job boards, the process of headhunting or targeting job adverts at individuals through social media is often done via the same algorithmic profiling used in all modern digital advertising, meaning it can be prone to some unexpected biases.

In one study, adverts promoting jobs in science, technology, engineering and maths were placed on Facebook. The results found that the ads were less likely to be shown to women than to men, even though neither the researchers placing them nor the Facebook algorithm itself were aiming to recruit more men for the posts.

The explanation the researchers found was that because young women are a valuable demographic for Facebook ads in general (because they typically control a high share of household spending), ads targeting them are therefore more expensive than ads shown to young men. The algorithms, by trying to maximize cost-effectiveness of the advertising spend, therefore targeted the ads more towards men than women. (Lambrech and Tucker, 2018)

CV screening is the biggest area of algorithmic use in recruitment at present, where it is mostly used to help reduce the time required to filter the very large volume of CVs that many job postings currently attract. These algorithms search CVs for keywords indicating essential skills or qualifications the company is looking for. Simpler algorithms might be programmed directly by humans as to which words or phrases to seek out, while more advanced machine learning algorithms learn for themselves what makes for an attractive CV based on comparing them to large sets of training data.

Victoria MacLean, a former recruiter for the finance sector, is the founder of City CV, a company that helps job applicants with preparing their applications. She noted that applicant-tracking systems (ATS) "reject up to 75% of CVs, or resumes, before a human sees them" (The Economist, 2018).

One potential benefit of these algorithms is that they offer the opportunity to reduce the unconscious bias that afflicts human

recruiters. Experiments from over 15 years ago, long before these algorithms were in common use, found that human recruiters screening CVs exhibit strong biases – in some cases 'white-sounding names' on CVs received 50% more interview offers than 'black-sounding names'; a bias worth as much as an additional eight years of work experience (Bertrand and Mullainathan, 2004). Machine learning algorithms can be used to help companies identify where biases might exist in their recruitment processes. Algorithms can also be used to carry out redacting of gender or race identifying information on CVs before they are seen by humans (or other algorithms).

However, when used clumsily, recruitment algorithms can perpetuate and even expand on human biases. The most well-known example is Amazon, which had been working since 2014 to build its own automated CV screening algorithm. They trained the algorithm using their own recruitment data going back 10 years, to learn what the company valued in its recruits. However, they quickly found it had learned unfortunate lessons from the fact that the company had previously more men than women. "In effect, Amazon's system taught itself that male candidates were preferable. It penalized résumés that included the word 'women's', as in 'women's chess club captain'. And it downgraded graduates of 2 all-women's colleges, according to people familiar with the matter." (Dastin, 2018)

While Amazon was able to remove the ability of its algorithm to rely on those particular terms, they found it hard to be sure they'd purged all ability of the system to discriminate. In effect, the algorithm was tainted from the get-go by the biases of previous human managers in the original recruitment data it was trained on. Amazon eventually abandoned the project by early 2017 after they lost hope in ever correcting the problems and finding a host of other issues, such as the algorithm recommending unqualified people for random jobs.

The main lesson from the Amazon case is that training an algorithm to learn from your historical approach to recruitment (or any other issue) will invariably lead to bias and discrimination insofar as the organisation's history and culture already exhibits bias and discrimination. The careful selection of training data is therefore crucial if algorithms are to realise their potential of removing bias – a point explored in more depth in the dedicated section on bias below.

Expanding on the roles of algorithms in recruitment, a few firms are now adopting algorithmic approaches to job interviews themselves. Judging the suitability of candidates in an interview is typically a complex process that can rely on difficult assessments about people's skills, character and how they would fit with the culture of the organisation.

However, companies like HireView have been making a name for themselves recently by promoting their technology that conducts remote video interviews with candidates, without a human interviewer, recording the candidates responses and then using AI to analyse their answers including their vocal and facial cues to give an assessment of whether they should progress to the next stage of the recruitment process. The use of HireView AI has been reported at large companies like Vodafone and Intel as well as major banks like Goldman Sachs and JP Morgan.

Automated approaches are also being taken to deep background checks for prospective employees; tools from companies like Checkr (used to vet Uber and Lyft drivers) or Fama Technologies can scan the social media feeds of prospective hires to check for signs of racism, misogyny or other potentially offensive language, helping to prevent potentially problematic hires.

A final use of algorithms in recruitment is through the automatic filtering of candidates through online assessments or psychometric tests. Ultimately in these cases the algorithm is only as effective as the test itself. Online, automatically scored assessments can be extremely useful in testing either for specific skillsets in an impartial manner, or in testing for general problem solving and analytical abilities, such as the online reasoning tests that have historically been used by the UK Civil Service Fast Stream and many other organisations.

Many firms may opt for their own in-house testing, but there is growing use of third-party personality testing through firms like [Pure Matching](#), who claim their algorithm "maps your neuro-personality as to gain an overall picture of your biological identity. This allows

us to map out your personality as well as the person you are being matched with in great detail and bypass the pitfall of matching by means of skills on a CV only."

Overall there is little doubt that the use of algorithms can be a major help to recruiters, not least in speeding up the process of filtering very large numbers of applications. At Unilever, for example, the average recruitment time has been cut by 75% thanks to the use of automated screening processes (Heric, 2018). German company SAP conducted a review of their new recruitment process involving algorithmic CV screening, 2 algorithmically assessed online tests and only then passing shortlisted candidates to a human to interview.

They concluded that:

"The percentage of abandoned applications dropped from 93% to 25%, meaning that a far higher number of valuable candidates continued with the application process and ultimately got hired. The company's cost savings on recruitment alone was projected at over £250,000 in the very first year following the algorithm's implementation. SAP received no complaints about the process. On the contrary, the graduates by and large rated the online tools highly, as 75% said that they increased their motivation to apply, and 88% claimed to have been more engaged with the process than with others they had encountered." (Hopping, 2015)

While these tools can be very helpful in complementing human interviews and helping to counteract human biases, there is a danger if too much of the recruitment process is turned over to machines. 61% of new job applicants would prefer face-to-face interviews to digital recruitment methods, according to a survey by ManpowerGroup Solutions. 42% of applicants meanwhile say that technology dehumanises the recruitment processes and worries that it might screen in or out the wrong people. Generally though, a majority of candidates who had interaction with algorithms in the recruitment process, in the form of chatbots, CV reading software or automated testing, reported a positive experience. (TribePad, 2019)

Some firms are using recruitment algorithms precisely to improve the experience for applicants; for example, Mya is a chatbot algorithm employing natural language processing to engage with applicants throughout an ongoing recruitment processes, answering their questions and keeping them informed as their application progresses – the software is in use at major firms from PepsiCo to L'Oréal.

The foundations of the EU's General Data Protection Regulation (GDPR), set out at the beginning of the document, include a reference in paragraph 71 to "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'."

Companies need to be careful that whatever algorithmic approaches they are using, it does not fall foul of this provision. Media reports such as the following case suggest that some companies might be rejecting applicants based on automated results, without their applications ever coming to a human's attention:

"Harry, 24, has been searching for a job for 4 months. In retail 'just about every job opening' requires a test or game. He completes 4 or 5 a week. The rejections are often instant, piling up without a word of feedback. Every time you start again from zero. 'You never know what you've done wrong. It leaves you feeling a bit trapped,' Harry says." (Buranyi, 2018)

Cathy O'Neil highlights another example of a young American man with bipolar disorder, Kyle Behm, who was repeatedly rejected from jobs because of their use of automated 5 factor personality testing that was screening out candidates who scored high on neuroticism; resulting in a lawsuit in the US over whether firms were illegally carrying out medical exams and discriminating against those with mental health problems (O'Neil, 2016).

Similarly, use of technology that analyses facial cues such as HireView might be penalising applicants with autism or other conditions that affect their facial expressions. Facial recognition AI has often been criticised for failing to properly identify or read non-white faces, leading some people to question whether the use of facial-analysis AI in recruitment processes poses a similar risk of bias. In the

USA, a privacy watchdog has asked the Federal Trade Commission to investigate Hireview for 'unfair and deceptive practices', on the grounds that its facial analysis and algorithmic decision making is not transparent. Any use of such technology in the UK should be carefully checked against the requirements of the Equalities Act.

Another danger is in inadvertently giving algorithms more weight in the process than is intended. When humans still make the final decision, evidence suggests they may be disproportionately affected by the ranking given to candidates by an algorithm, picking the 'number one' candidate even if their score was only insignificantly different from the 'number 2' or even any other of the top 5, rather than using their human judgement to fully evaluate all the candidates the algorithm considered suitable. "Hiring tools that assess, score, and rank jobseekers can overstate marginal or unimportant distinctions between similarly qualified candidates. In particular, rank-ordered lists and numerical scores may influence recruiters more than we realize, and not enough is known about how human recruiters act on predictive tools' guidance." (Bogen and Rieke, 2018)

A final problem with recruitment algorithms is that their widespread use encourages candidates to try to game the system. According to one survey, 88% of applicants who know about algorithmic recruitment systems have 'optimised' their CVs as a result and such applicants are 4 times more likely to say they have 'cheated' on a test. Companies like Practice Aptitude Tests have sprung up precisely to help coach people through these kinds of algorithmic assessments, while Victoria McClean founded City CV to help candidates deal with recruitment algorithms, recommending things such as preparing different wordings of the CV for different jobs and copying acronyms used in the job description in case those are what the algorithm is searching for.

A much more positive use of AI in recruitment comes from start-up firm Textio. Their machine learning software analyses job listings and assesses to what extent certain words and phrases attract or put off potential applicants, including on grounds of bias. By flagging words such as 'expert' or 'aggressive' in job adverts as conveying an overly masculine tone, they can be reworded to more gender-neutral phrases in a way that is likely to attract a larger and more diverse set of applicants.

The conclusion appears to be that algorithms can be used to improve recruitment quality and reduce bias, or they can be used to reduce recruitment time and cost and improve efficiency. While not entirely mutually exclusive, these 2 aims are often in tension and firms need to think carefully about why they are using such algorithms and what they hope to achieve before adopting them.

Use of algorithms in employee management

One of the main ways in which algorithms are both assisting and, in some areas, even replacing human managers is in work allocation; either in allocating workers to particular shifts or teams or in allocating tasks to individual workers. Shift allocation algorithms can automatically match workers to shifts when they are needed and also handle the routine business of allowing workers to swap or change their shifts.

Increasingly common in the retail and hospitality sectors, they can also include quite sophisticated machine learning algorithms to forecast customer footfall, using anything from traffic history and point of sale data to weather forecasts. These predictions are then used to match to employees' skill sets and calculate which employees should be scheduled on any given day, in order for workers' shift patterns to respond to consumer demand. Platforms like Rotageek are in use by companies including Pret A Manger, O2 and Thorpe Park while Percolata is being employed at UNIQLO.

These abilities to generate shift schedules in response to changing demand obviously can benefit firms and consumers by making sure people are working at the times when they are needed. They can benefit workers by giving clearer advance notice of when shifts will be and making it easier to swap and change them. This offers a potentially major benefit to workers in industries like retail or hospitality, who at present face often being on call or ready to turn up for shifts, only to find them cancelled or cut short with very little notice. Having to go through a human manager to swap a shift with another willing colleague is also an added burden that can often put workers off taking more control of their schedules.

Automating shift scheduling also potentially removes the ability of line managers to exercise personal favouritism and bias in the allocation of shifts, though in practice these platforms report that some managers deliberately avoid using the auto-scheduling functions precisely for this reason, letting the algorithms advise them on recommended shifts but keeping the option to overrule them in order to retain a level of personal control over their workers. As Percolata founder Greg Tanaka has commented, "What's ironic is we're not automating the sales associates' jobs per se, but we're automating the manager's job, and [our algorithm] can actually do it better than them." (O'Connor, 2016)

Concerns have been raised over whether such tech could be used in a less benign manner to control workers on zero-hours contracts, giving them just enough shifts to keep them with the organisation, or auto-scheduling shifts so that all workers are constantly kept just below the legal thresholds for employee status, in order to deprive them of full employment rights. Such algorithms also typically allocate more shifts to workers it calculates are the highest performing – while preventing shifts being allocated based on a managers' favouritism, this still can put a lot of psychological pressure on some workers towards trying to please the algorithm, distorting workers' behaviour towards whatever it is that the algorithm happens to be measuring.

Far beyond just matching workers to shifts, however, algorithms can replace much of the day-to-day allocation of tasks to workers that has traditionally been done by human managers. The in-house Preactor software in use at Siemens manufacturing plants can plan production orders in real time, saving time and increasing flexibility, but reducing the autonomy of individual factory workers in the selection and ordering of their day-to-day tasks, posing a dilemma as to how much managers should allow deference to such software to replace human judgement. "I think that you have to manage that bit very carefully because if someone said to me that 'you don't have to think anymore, you just have to do whatever the screen tells you', I'd find that really hard." (Briône, 2017)

Some logistics firms take this kind of task allocating to significant lengths. Delivery drivers for some firms often spend their entire working days following algorithmically generated instructions on what route to take to their destination, in order to fulfil jobs that are themselves being allocated to them by an algorithm, by an algorithmically calculated target time that they are under huge pressure to beat, or see their performance downgraded by yet another algorithm. Regardless of the classification of their employment status, such workers have very little contact with human managers in the course of their jobs.

The same can be true of warehouse workers; handheld devices and tablets have long been used to give warehouse 'pickers' sets of timed instructions as to what items to collect from where on a minute-by-minute basis. Amazon warehouses are now taking this to the next level – workers are being equipped with a wearable haptic feedback device that tells them what to collect, where to find it in the warehouse and gives them a requisite number of seconds to find the item. They wear these devices on their arms and it uses vibrations to guide their arm movements in order to be more efficient.

Taking away people's autonomy in this way can remove an important sense of dignity and humanity from work, when workers are denied the ability to make even tiny or mundane decisions about what size of box to use or how long a piece of tape to cut for wrapping, or even where and how to move their own limbs.

Algorithms also take away control from platform workers; they are not told how long the next trip will take for example or where it will leave them before they click to accept it. Some delivery couriers have historically not been told the delivery destination address until after they picked up the package or food from the restaurant – by which point it was too late to turn it down without the algorithm rating down the courier.

Uber and Lyft drivers complain that the algorithms used to allocate their jobs make it difficult to turn them down, even in subtle ways, "when they show the spot on the map where you're going to pick someone up its very zoomed in so if you're not immediately familiar with the area you probably wouldn't be able to discern within 12 seconds if its somewhere you want to go or not. They just tell you how far away it is in driving time." (Lee, Kusbit, Metsky, and Dabbish, 2015)

It is when these gig economy platforms combine with the consumer demand-forecasting algorithms that we verge into the gamification area of targeted incentives discussed earlier, that borders on psychological manipulation. The algorithms end up using gamification to nudge people to work during hours they would normally prefer to spend not working through the carefully calculated use of personally targeted micro-incentives.

There is growing evidence that some of these digital management practices pioneered in the gig economy are starting to spread to the wider labour force. A 2019 study estimated that over 12 million workers in the UK in total have been having their work logged through digital apps or websites and that "for every platform worker who has used an 'app' or website to log work, there are 1.2 non-platform workers who have done so." At the same time, the authors noted "the sharp rise in the use of digital means (apps or websites) for notifying workers when new tasks are waiting for them. In 2016 one person in ten was reporting this practice, but by 2019 this had more than doubled to 21.0% of the adult working-age population." (University of Hertfordshire, 2019)

The lack of contact with human managers also has obvious implications for the wellbeing of workers, but also for the performance of their jobs – research suggests that workers cooperate much less with work instructions given to them by machines rather than people. (Parise, Kiesler, Sproull and Waters, 1999)

Workers also seek to find ways to manipulate such algorithms, finding workarounds to compensate for their inflexible rules. For example one study of rideshare drivers found that "when drivers desired a break but did not want to turn off their driver applications to benefit from an hourly payment promotion, they parked in between the other ridesharing cars in order not to get any requests," as they knew the algorithm would allocate the closest driver and therefore felt safe if they were between 2 other vehicles.

More transparency in the way in which such algorithms assigned tasks was particularly helpful to workers. The paper concluded that "our findings suggest drivers benefited from deeper knowledge of the assignment algorithm. Drivers with more knowledge created workarounds to avoid undesirable assignments, whereas those with less knowledge rejected undesirable assignments, lowering their acceptance rating, or unwillingly fulfilled the uneconomical rides." (Lee, Kusbit, Metsky and Dabbish, 2015)

Work allocation algorithms can be used for far more radical approaches to workplace management, however, rather than simply being tools for micro-managing daily tasks. Some firms have been experimenting with revolutionary approaches focused on a complete redesign and restructure of work teams and departments, enabling agile working on a new scale.

"Publicis, a multinational marketing company, has already started using algorithms to organize and assign its 80,000 employees, including account managers, coders, graphic designers, and copywriters. Whenever there is a new project or client pitch, the algorithm recommends the right combination of talent for the best possible result." Publicis shift around their workers regularly as the algorithms recommend, every time a project is started or finished. (Walsh, 2019)

For managers themselves there is a conundrum; these algorithms allow them potentially much greater control over their workforce but at the cost of paradoxically making their own jobs less relevant; if all major recruitment, task-allocation and performance review functions can be undertaken by algorithms, what discretion is there left any more for human line managers?

In the long run such algorithms could threaten the existence of human line managers altogether – something already seen in gig economy firms where delivery drivers or warehouse workers receive their day-to-day instructions directly from a tablet or device without the need for a human manager. This might explain why Percolata reports that the use of its auto-scheduling function for store shifts so often goes unused by store managers at its client companies, who prefer to retain the discretion to make their own decisions even while taking advice from the algorithm on projected footfall.

Use of algorithms in performance management

The third broad area of algorithmic management lies in monitoring and assessing the performance and behaviour of existing employees. A lot of this technology revolves around data collection; whether from customers, from employees willingly entering survey data, or through automated tracking and monitoring software that collects data on employees' activity.

Perhaps the most archetypal use of this technology tends to be found in call centres, which have easily measurable metrics for performance and productivity. One such tool call centres have been using is Cogito, an AI programme that both provides live and recorded activity and productivity assessments of each worker to managers, but also provides individual employees real time feedback on their performance based on more advanced voice analysis of their conversations: "Talking too fast? The program flashes an icon of a speedometer, indicating that he should slow down. Sound sleepy? The software displays an 'energy cue,' with a picture of a coffee cup. Not empathetic enough? A heart icon pops up." (Roose, 2019)

This is a new generation of performance management software that does far more than just log hours or record activity – it can approximate in some narrow areas (like making phone calls in a call-centre) the same kind of detailed feedback that workers might receive from a human line manager.

More generally applicable across different industries, many firms are adopting technology that uses Microsoft's LinkedIn to regularly survey their employees and detect any changes in performance or morale. LinkedIn also launched a product in September allowing employees to check in their own performance goals and compare themselves to company-wide benchmarks, allowing a form of self-service performance review.

The use of customer ratings to assess employees' performance in aggregated scores is another practice that is gaining ground, seeping over from the gig economy to the regular workforce. According to a 2019 workforce survey, over half the UK working age population reported having been subject to customer ratings at some point. (University of Hertfordshire, 2019)

Some companies, however, are able to draw together streams of data from a whole host of different sources, using custom platforms. At IBM, they have been using their Watson supercomputer to predict employee attrition and take steps to prevent it.

As they describe:

"Using Watson algorithms, the HR team developed and patented a program that looks at patterns in data from all over IBM and predicts which employees are most likely to quit in the near future. The algorithms then recommend actions—like more training or awarding an overdue promotion—to keep them from leaving."

While their managers don't have to follow the recommendations of the algorithms, in practice the company has been able to convince most of them of its benefits. "At one point, all the data showed that giving a certain group of employees a 10% raise would reduce their 'flight risk' by 90%," commented Diane Gherson, IBM's chief human resources officer. "Managers who didn't take that advice had attrition rates on their teams that were twice as high as for those who did."

Another way IBM won over sceptical managers, she adds, was by "explaining why the system recommends a certain action. You have to open up the black box a bit and show people the data." (Fisher, 2019)

Algorithmic assessments can also be used in a more formal performance review role - either to replace human manager-led performance reviews, or to support them by adding automated assessment data for managers to use as a basis of discussion in appraisal meetings. One of the benefits of automated performance review include that it can be done in real time, rather than in annual appraisals, and can correct for a number of human biases human managers exhibit:

"Not only can AI correct for racial and gender bias, but it also is not susceptible to performance-review-specific biases, such as recency bias (where actions performed recently are given more weight than actions that occurred say, 11 months ago for a yearly

assessment). Similarly, AI can control for contrast bias, which occurs when a manager compares an employee's performance to their peers rather than to objective measures of success. This bias can be particularly pervasive in growing companies - perhaps the entire sales team met their goals, but an evaluator may be inclined to give the least successful representative a worse performance review, even though they objectively performed to standard." (Eder, 2018)

However, as with other uses of algorithms, such elimination of biases can only be achieved based on a very careful selection of training data and rigorous approaches to removing unnecessary demographic cues and careful auditing of the algorithm. Another concern that has been raised is whether algorithmic performance review pushes workers towards hitting targets that are assessed by the algorithm, rather than simply trying to do a good job. One paper claims that "algorithm-based HR decision-making can shift the delicate balance between employees' personal integrity and compliance more toward the compliance side because it may evoke blind trust in processes and rules, which may ultimately marginalize human sense-making as part of the decision-making processes." (Leicht-Deobald, and others, 2019)

Such performance review and HR algorithms are also frequently reliant on heavy monitoring and data collection about employees, in order to provide them with the information they need to make assessments. These kinds of data collection for algorithmic purposes can cause problems in and of themselves, quite apart from whatever decisions might ultimately be made of the back of them.

The simplest kind of monitoring algorithms are activity monitors, which for example record how long employees spend at their desks – OccupEye being one such technology seeing use at UK firms such as the *Telegraph*, where it was introduced with no warning or consent by the workforce, leading to a massive backlash from the National Union of Journalists and a hasty U-turn by management. These kinds of monitoring can lead to presenteeism pressure. An ethnographic study of long-distance haulage drivers showed, for instance, that electronic monitoring led them to feel pressured not to take their mandated rest breaks. (Levy, 2015)

More advanced products such as Humanyze, Isaak or Plasticity Labs instead attempt to analyse more complex metrics such as workers' mood or mental state, as well as their interactions with others. Humanyze for example "is a credit card-sized device worn by workers to monitor their mood and understand team dynamics, such as who is usually dominant in conversations and who appears most engaged. It draws on infrared sensors, microphones and an accelerometer." (Dellot, 2017)

This offers the potential for a more useful and human-centred understanding of the workforce, but might also be considered to be even more intrusive by workers.

It is in use in the UK at Deloitte (management consultant), a high street bank, retailers and some parts of the NHS. "Workers in these trials did appear 'enthusiastic'. None of the employees were forced to wear the device, but the company claimed 90% opted to do so." However, Pam Cowburn, communications director at Opens Rights Group, commented that "Staff may feel [pressured] into consenting to wearing surveillance devices because they fear that they will be discriminated against if they don't." (Kobie, 2017)

The Isaak system designed by London company Status Today meanwhile is similarly being used to analyse how much workers collaborate, classifying key workers them as 'influencers' or 'change-makers' by combining activity monitoring data with data from sales performance figures or personnel files. It is currently in use by several law firms, estate agents and other companies totalling 130,000 UK employees as of early 2019. Status Today argue that the data could be used to help better share responsibilities between workers, "ultimately improving the overall workplace environment and reducing stress and overworking," though their chief executive admits that "there's always a risk that it might be misused" to focus only on increasing productivity without addressing employee wellbeing. (Booth, 2019)

According to Chris Brauer, director of innovation at the Institute of Management Studies at Goldsmiths, "there are countless advantages to these kinds of technologies in the workplace... They help make visible analytics and data around people instead of just looking only at machines or process flows that can really tend to dehumanise a workforce. You can identify pockets of innovation that need support, clogs and limitations to communications, promote collaboration and teamwork, and design and orient physical space in

the interest of a healthier and more productive work environment." (Booth, 2019)

AI can also be used to give managers a better understanding of their direct reports and how to interact with them as individuals, instead of using algorithms to treat them as interchangeable cogs in machine. At Cisco, UK vice president Eleanor Cavanagh-Lomas told *Fortune* magazine back in 2016 how she used an in-house people analytics app called Team Space to see an algorithmic assessment of her 15 direct reports based on tests they all took, allowing her to see recommendations of how each person works best and how to approach them.

"When Cavanagh-Lomas learned that this manager doesn't give up easily, she acknowledged that quality in a conversation and asked him to try a different tack for a short time. If it failed, he could return to his method again. 'The software gives you coaching tips tailored to their own style and how they need to hear feedback,' she says. It worked, she adds." (Alsever, 2016)

Overall the use of AI as part of HR offers many opportunities to better understand the workforce, improve on staff retention and carry out improved staff performance assessments. There are dangers, however, in simply outsourcing performance management decisions that could have big effects on employees' pay and career prospects to an algorithm.

An ILO report into the use of such technology raised concerns that "A risk of PPVH [physical and psychosocial violence and harassment] arises when data is used to make seemingly neutral decisions about performance, and when targets are universalized and do not take into account, for example, physical differences between workers, some of whom may not be able to work faster for health reasons." (Moore, 2018)

To combat these risks, it is important both that human managers remain closely involved in the process, and that workplace ethics are at the forefront of the design of all HR algorithms. According to an EU report, "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." (European Agency for Safety and Health at Work, 2019)

Again, it comes down to how the technology is used. A responsible and positive use of performance management algorithms would be using AI to help identify under-performing workers for the purposes of providing additional support, or working out what additional factors might be contributing to the under-performance and working to address them. A less responsible use would be to simply identify the bottom 10% of performers and sack them; an approach that has been anecdotally reported as prevalent in workplaces like distribution warehouses or call centres where employee turnover is high and work is often short-term and casual.

Opportunities and risks

Productivity

As Dr Phoebe Moore put it, "The reasons algorithms are interesting is it creates the idea there are ways to make decisions that are more efficient than what humans can do alone". Certain tasks that a human statistician might take a year to do could be done in 5 minutes by a powerful enough computer.

The main attraction that algorithmic management poses for many employers is the speed at which it can process certain HR tasks compared with human managers. Time saved equals productivity gained and in an area of stagnant productivity in the UK it is not surprising that employers will be looking for every edge they can get. The time of human line managers is particularly valuable and the less time they have to spend on routine, automatable tasks such as allocating daily tasks and drawing up shift rotas, the more time they can (in theory) devote to areas where a human being has real advantages over an algorithm; coaching, supporting and

developing their workers, and resolving disputes in the workplace, all of which can impact on productivity.

Ranking, sorting and filtering tasks are particularly well suited to automation; in this sense recruitment is probably the area where algorithms can have the greatest impact on productivity. In a globalised world and with high degrees of labour mobility, some companies may now receive hundreds or in some cases thousands of job applications for each post, from all around the world. In these situations, it simply is not feasible to sift through all the initial applications by hand. Providing issues of bias and accuracy can be tackled, turning over CV screening to algorithms makes perfect business sense, even if a human manager should have ultimate oversight and still lead on the final interviewing and decision-making.

The use of algorithms in task allocation can also not only save management time, but lead to much more efficient decisions, allocating and reallocating tasks in response to changing demand, to make sure that workers are always assigned to the most useful thing they could be doing given the current circumstances – something particularly vital in manufacturing firms that rely on constant just-in-time supply chains. The use of the Preactor software at Siemens was described as "enabling production orders to be sequenced and planned to the point of everything working in a way that stops bottlenecks, in a way that makes sure that if we have a delivery date we are meeting it and the flow through the factory should be seamless." It even enabled managers to update the complex build plans for the factory floor in real time, something that was near impossible when relying on human managers to develop the production plans. (Brîône, 2017)

Performance review algorithms meanwhile, as discussed above, can be used to help identify certain factors that may be contributing to underperformance, providing targeted support to help struggling employees improve their productivity, or providing the right targeted incentive structures at the right time to get the best out of a company's workers.

In some cases, the savings from automated management, rather than freeing up human managers to focus on other things, are used to eliminate the role of human managers altogether. In the case of many gig economy firms, the platform workers are managed entirely by the software platform, receiving tasks, instructions and being rated on their performance, without a human manager being involved at all.

While this poses serious questions about the quality of work and dangers of dehumanisation, there is no doubt that the primary motive for this model of working is cost and efficiency; the entire business model of these firms relies on not having to employ human line managers to interface between customers and workers. Whether in the long run this is a positive and sustainable business model is another question.

Can decisions without humans still be humane

It is worth noting that Section 4 of the GDPR outlines a 'Right to object to automated individual decision making'.

Article 22 on 'Automated individual decision-making, including profiling', states 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."

Such processing includes "any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements."

[Article 4(4)]

However, the UK government has exercised its right under Article 22(2).b of the GDPR "to obviate users' right to not be subject to automated decision-making" (Access Now, 2019). In the UK automated decision making therefore is allowed, but subject to various safeguards; a right to receive notification of automated decision-making, which must be provided "as soon as reasonably expectable";

a right to contest automated decisions and demand human intervention or reconsideration; and, in response to this contestation to receive information on what steps were taken to comply and the final outcome of the decision.

According to the ICO's published guidance in the UK:

"You can only carry out solely automated decision-making with legal or similarly significant effects if the decision is:

- necessary for entering into or performance of a contract between an organisation and the individual
- authorised by law (for example, for the purposes of fraud or tax evasion)
- based on the individual's explicit consent.

"If you're using special category personal data you can only carry out processing described in Article 22(1) if:

- you have the individual's explicit consent
- the processing is necessary for reasons of substantial public interest" (Information Commissioner's Office, 2019)

Awareness of even the limited application of GDPR to automated decision making in the UK appears to be extremely low. There have been no ICO imposed fines recorded under this section of GDPR in the UK. Concerns have also been expressed that, in part due to low awareness, many UK companies adopting algorithmic management tools are buying off the shelf packages from companies based in the United States which has much looser regulation on matters of surveillance, data protection and consent. Some of these tools may not in fact be legally suitable for a UK workplace environment.

As Dr Phoebe Moore describes it, many start-up companies in particular "want to experiment with their workforce but they're not looking at the legislation – employment, social, data protection, because they want to power ahead."

One of the issues with this section of regulation is that it only applies when decisions are entirely automated; a limited degree of human intervention in the process could still be enough to render the requirements to notify and allow contestation as not applicable. This human intervention could be a human reviewing a list of recommendations from an algorithm and then clicking 'approve', though in such cases it might be hard to know whether the human was really actively reviewing them at all or just passively accepting all the algorithm's recommendations.

Nonetheless, in cases where responses are almost instantaneous, such as an immediate email back from a job application saying you have been unsuccessful within seconds of applying, it is fairly clear that the decision was not even subject to this limited level of human oversight but was entirely automated. The fact that such outcomes are being routinely reported in the press suggests that there may be serious problems with compliance in this area currently going unreported and unchecked.

It's worth also looking at the increasing automation of certain HR functions, such as systems for approving annual leave. On the one hand, removing managers' discretion from these tasks could improve fairness by ensuring everyone is treated the same according to a pre-determined set of rules; removing the ability of bad managers to show personal favouritism or use the threat of withholding leave as a tool to pressure and control employees.

On the other hand, it also removes any element of human compassion from the process; employees who might have special circumstances that a human manager would recognise and grant extraordinary leave for would have a much harder time convincing an algorithm to make a special exception for them. It could be argued, then, that algorithmic management in these areas is better for workers than having a 'bad' line manager, but worse than a 'good' one.

As Prospect's submission to a consultation on the future of work put it: "New technology also raises questions around the skills managers have to understand the data collected on workers, and whether this may lead to institutionalising poor management practices. Management-by-algorithm will also cause challenges in analysing contextual factors, such as compassionate leave or the impact of stress at work." (Prospect, 2019)

As Professor Ursula Huws described it when interviewed for this research, "Interposing a digital interface between the worker and the line manager strips away a lot of the humane aspects of HR."

Where the task allocation and other 'routine' functions of management are outsourced to algorithms, it's important to look at what happens to line managers. In some firms this enables line managers to focus more on the interpersonal aspects of their jobs, developing and supporting their direct reports. In other cases, however, stripping out many of the day-to-day functions could lead to the role of line managers more generally being undermined and the pastoral and mentoring roles which they may have previously had either being lost, or being diluted as they are now allocated far more employees per manager due to their increased capacity as a result of the algorithm assisting them.

Professor Huws points particularly to the replacement of recruitment, training and broader onboarding of new employees with automated systems as undermining the early settling in of new employees to the organisation and introduction to its culture which is vital to their longer-term engagement and career development. When your first introduction to a company is all through automated systems, it can set the tone for the future of the employment relationship and make it harder to get a good feel for the organisation's culture and values. It can also be inherently dehumanising for workers to feel like they are being directed by a machine rather than another person.

On the other hand, many of the most promising growth areas for AI in HR make use of algorithms to support and assist decisions that are still ultimately being made by human managers. By providing managers with more data about the workforce, it can help them to understand their workers better as individual employees, understand their needs and behaviours and how they interact with one another.

Automated self-service HR tools for things like booking shifts and annual leave can also greatly improve convenience for employees – providing that there is still an option to talk to a human manager who can make the needed exceptions for workers when their personal circumstances warrant it. In these ways algorithmic tools can be used to improve the quality of work for their employees – but only when they are used to support decision making by human managers, rather than replace human managers entirely.

Bias

One of the biggest areas of concern in the use of algorithms is evidence that algorithms in practice can embed and perpetuate biases and discrimination, often in difficult to detect ways. Examples of such biases have been widely reported in the press over the past few years, from the automatic soap dispenser that didn't recognise non-white hands (source: [Daily Mail article](#)) to accusations of sexism in the credit scoring algorithms in the Apple Card (source: [BBC Business News](#)), plus the gender bias in Amazon's automated recruitment experiment discussed in more detail above.

Experts on the subject are clear, however, that the main reason algorithms display biases is because of the biases of the humans they learn from. A return to analogue human decision making will not be unbiased either – in many cases even when algorithms demonstrate bias they can still perform better than human decision makers, something repeated studies have shown to be possible. (Miler, 2018)

To quote the title of a recent article by Stephen Bush in the New Statesman, "Of course algorithms are racist. They're made by people." However, as Bush argues, "even the algorithms that produce incorrect and prejudiced results more than half the time perform

better than the humans they have replaced. To take the example of the Metropolitan Police's use of facial recognition tech: just 2 out of the 10 people it flagged were flagged correctly ... but is quite literally 100% more effective than the old powers of stop and search, where just 1 in 10 stops ended in an arrest." (Bush, 2019)

We shouldn't, therefore, rush to the simplistic conclusion that algorithms are necessarily going to make bias worse. In fact they have the potential, when used wisely, to significantly improve on human decision making. We should never fall prey to the dangerous assumption that algorithmic decisions are truly impartial or 'objective'. Indeed, do they need to be? Don't they just have to be less biased than we are?

The last few decades of psychology research have made clear that unconscious bias is widespread and seriously affects a large part of human decision making. The study where Anglo-Saxon sounding names on CVs received 50% more interview offers than black-sounding names is an example that shows just how low the bar of human decision making is set that algorithms need to beat. (Bertrand and Mullainathan, 2004)

To make less biased algorithms, though, we first have to acknowledge and grapple with our existing human biases. Daniel Kahneman, Nobel laureate and one of the world's leading expert on biases, expressed the problem succinctly in his interview with Erik Brynjolfsson: "In the example of sexist hiring, if you use a system that is predicatively accurate, you are going to penalize women because, in fact, they are penalized by the organization. The problem is really not the selection, it's the organization. So something has to be done to make the organization less sexist. And then, as part of doing that, you would want to train your algorithm. But you certainly wouldn't want just to train the algorithm and keep the organization as it is." (Brynjolfsson, 2018)

The most important factor in reducing bias in algorithms is this need for good training data sets. "[A]n algorithm is only as good as the data it works with," according to Solon Barocas and Andrew Selbst. "Even in situations where data miners are extremely careful, they can still affect discriminatory results with models that, quite unintentionally, pick out proxy variables for protected classes." (Barocas and Selbst, 2016)

Hence attempts to avoid feeding recruitment algorithms with information about the gender or race of applicants can be undermined if the algorithms find proxies for such data elsewhere in the CVs such as the fact that an applicant attended a girl's school or a school in a predominantly ethnic minority neighbourhood.

A related consideration is therefore to avoid using too many variables in the datasets. The more data that machine learning algorithms have to work with, the higher the probability that they find something that is a good proxy for a protected characteristic that they could end up discriminating against.

Nevertheless, even when the training data is carefully selected, algorithms can still pose risks of discrimination according to the World Economic Forum (WEF) Global Future Council on Human Rights and Technology (WEF, 2018).

Their 2018 report points to particular dangers in cases of:

- Choosing the wrong model for the job
- Building a model which has inadvertently discriminatory features
- The lack of proper human oversight and involvement
- Unpredictable and inscrutable 'black box' systems that are hard to understand
- Unchecked and intentional discrimination decisions by humans in the process

The first of these might include examples of things like the technology analysing facial recognition cues in interviews, where the evidence of relevance and effectiveness is pretty lacking.

The second kind of error is typified by the recruitment algorithms that are trained on a dataset that contains examples of real-world

gender discrimination going back many decades.

The third example would refer to cases that would fall within the scope of GDPR regulation 22(1), where decisions are being made that can't be reviewed or challenged by human managers.

The fourth danger relates to cases where it's impossible to give an explanation as to why a certain decision has been reached, such as a decision by an algorithm to recommend a worker for recruitment, promotion, shift allocation or pay rise (or not to do so). Even if such decisions can be overturned on appeal, it's harder to make a case for doing so when it's not understood by either the subject or the reviewing manager how the decision was arrived at in the first place.

The fifth and final kind of problem is when the humans designing the algorithms do have explicitly discriminatory aims, such as to avoid hiring women likely to become pregnant, but are able to disguise such intentions through a complex algorithm that gives them the same end result but is more difficult to challenge in a court or tribunal due to the way in which the automated decision-making obfuscates the real rationale behind it. This may be the most pernicious threat of all.

This final kind of discrimination could also include discrimination introduced by consumers through ratings systems. An academic study of Uber looked at how customer ratings systems can introduce a "backdoor to employment discrimination." They found that the lack of information on how particular ratings related to specific behaviours was a major source of anxiety for drivers and that such ratings systems potentially created a way in which "companies may perpetuate bias without being liable for it." (Rosenblat, Levy, Barocas and Hwang, 2016)

Proving such discrimination may be extremely difficult in many cases, posing a real challenge for the ability of employment tribunals to provide redress. As a recent article in The Atlantic described, talking about accusations made against Facebook over biased news sources:

"Emphasizing algorithms over human actors can be very useful cover. While critiquing algorithmic practices in her book *If ... Then*, the technology researcher Taina Bucher builds on the social theorist Linsey McGoey's notion of 'strategic ignorance,' when it's useful for those in power to not know something. Rendering systems as wholly human or wholly algorithmic, Bucher writes, is a 'knowledge alibi,' a deflection wherein technology is conveniently described as independent and complex, but human actors as inert and oblivious to consequences."

Already Google, Apple and Amazon have proved very resistant to lawsuits that allege they privilege their own products in their search results or app stores, on the basis that the algorithms behind such rankings are extremely complex and hard to unravel.

However, there is one further way in which algorithms could prove a huge benefit in reducing bias. Instead of using algorithms to make workplace decisions, they can be used to assess decisions made by humans. Algorithmic tools can, for example, analyse company payrolls to measure the levels of gender or racial pay gaps in different parts of the organisation and what factors seem to contribute to them.

They can assess individual managers based on how often they recommend men versus women for recruitment, promotion or pay rises to identify those who might need additional unconscious bias training. They can also scan the content of internal communications or external job postings for gendered language terms and recommend alternatives. In this way algorithms could make a huge difference to eliminating the gender pay gap and other workplace disparities.

A final point worth considering is that in many situations it is not only difficult but mathematically impossible to eliminate all forms of bias completely. Part of the problem is conflicting definitions of what counts as fairness in the first place. Algorithmic mistakes come in 2 main varieties; false positives – the probability, for example, of a hiring algorithm incorrectly recommending recruiting a bad employee, and false negatives – incorrectly rejecting a candidate who would have been an excellent employee. Biases can manifest

in both varieties – falsely hiring more unsuitable men than women and/or falsely rejecting more suitable women than men.

An unbiased algorithm can also be assessed on its 'predictive parity' – whether the algorithm is equally good at predicting whether a male or female candidate is suitable, and 'demographic parity' – whether the same number of men and women are being hired overall.

Alternative measures include 'accuracy parity' or equality of opportunity – whether the same number of qualified men and women are being hired, and 'individual fairness' – whether any single individual is treated the same regardless of their gender.

While we might want to eliminate biases in all these areas, it has been shown mathematically that for any decision where the characteristic in question correlates at all with relevant factors, it is impossible to make decisions that are fair according to every one of these definitions of fairness.

“You can’t have it all. If you want to be fair in one way, you might necessarily be unfair in another definition that also sounds reasonable,” as Michael Veale, a researcher in responsible machine learning at University College London puts it (Courtland, 2018).

Improving fairness in one area might necessarily reduce it in another. A sophisticated approach to algorithmic management therefore needs to not only seek to reduce biases, but to have open and honest conversations about what kinds of fairness are most important in the first place.

Control and consent

A second broad area of concern relates to the use of algorithms by employers to coerce and control the workforce. According to Mike Walsh, author of 'The Algorithmic Leader: How to Be Smart When Machines are Smarter Than You', "without careful consideration, the algorithmic workplace of the future may end up as a data-driven dystopia. There are a million ways that algorithms in the hands of bad managers could do more harm than good." (Walsh, 2019)

In particular he points to the ways in which algorithms are tools that enable a new digital Taylorism, reminiscent of the management approach of the early 20th century with its heavy focus on quantification, monitoring, control and efficiency. Managers are generally happy to embrace technology that allows them to track and monitor workers to a greater degree, but line managers often don't like it when algorithms take away their own control or autonomy by making major decisions automatically.

The recent bestselling dystopian novel *Perfidious Albion* by Sam Byers depicts a near-future UK in which algorithms are used by powerful corporations to manipulate and control both workers and the general public, in insidious ways that are almost impossible to detect. While managers are using the software to manipulate and control frontline workers, they themselves are being manipulated and controlled using the same software by those above them in the corporate hierarchy. Some of the tools entering the market today make such a dystopian vision not that far-fetched.

PC Magazine in its review of top employee monitoring software discusses how algorithms can be used to enhance simple monitoring tools into a truly Orwellian situation, such that "if an enterprise's C-suite executives want to know whether employees are chatting internally about the company's CEO or CTO, they could simply set up automated keyword triggers to receive an email alert or have all mentions aggregated into a report." (Marvin, 2019)

Their top reviewed product in the category for 2019, Terramind, offers a list of features including 'stealth monitoring', 'live video feed', 'remote desktop control', 'document and file tracking', 'keyword tracking', 'optical character recognition', 'screenshots', 'automated alerts', 'keystroke recording' and 'location tracking', all accessible via a 'cloud dashboard'. The review concludes that "It's truly an all-seeing eye," though admits that "the depth of monitoring features can be daunting."

TUC research shows most UK workers (56%) now think it's likely their employer is monitoring them at work. This includes belief that their employer is using monitoring of work emails and browsing (49%), CCTV (45%), logging or recording phone calls (42%), handheld

or wearable location-tracking devices (23%) and facial recognition software (15%). The research also found, however, that most workers consider things like location-tracking, facial recognition, keystroke logging and other invasive forms of data collection to be unacceptable, while nearly 80% agreed that "employers should be legally required to consult and agree with workers any new form of workplace monitoring they are planning to introduce before they can enforce it. (TUC, 2018)

While not all forms of surveillance use algorithms (at its most basic, surveillance could just be a CCTV camera monitored by a human security guard), what algorithms do is allow the volume of surveillance data collection going on to grow far beyond the ability of human managers to monitor it all, by relying on filtering, matching and sorting algorithms to monitor the data instead and bring to managers' attention whenever something of interest occurs or arises.

This raises major questions around employee consent. In the UK employees have a legal right to be informed about surveillance or monitoring in the workplace – features such as 'stealth monitoring' that are built into platforms such as Terramind for their US customers are not lawful to use in UK workplaces. In the case of the more intrusive forms of monitoring algorithms which rely on collecting location tracking or other data from wearable tech, these usually require employees' consent.

There is a debate, however, around how much such consent is really informed or free in a context in which employees may feel pressure to take part – rather consent might be "manufactured", as Italian philosopher Antonio Gramsci would have described it, with people feeling pressured to consent by the power that their employer (or the tech company in other contexts) holds over them and the weight of cultural expectations from other colleagues signing up that normalises it.

This is not to say that surveillance can't be useful and important, or even for the benefit of workers. Microsoft, for example, has developed a 'smart camera' which uses AI to spot spillages, unmanned tools or other potential hazards in factories and warehouses. More broadly, better observation and analysis of workers' behaviour during their daily working lives could lead to identifying threats to their physical or mental health and developing better workplace design. Analysis of working time could be used not just to guard against workers' being underproductive, but also to watch for signs of overwork and help identify when workers are overstretched and need to be encouraged to go home and switch off. Again, the same technology can be used for very different purposes.

Major banks have also been rolling out new surveillance algorithms to guard against potentially fraudulent behaviour or other wrongdoing that could lead to a repeat of major scandals from the past decade. Similarly, where workers do have disciplinary procedures brought against them, or where they are bringing grievance procedures of their own, it could benefit both sides to have an objective digital record of what happened in the workplace; providing both sides are given equal access to such information when it would help their case. Workers bringing claims to employment tribunals could be greatly assisted by having data on their performance, activity and behaviour at work logged and securely retained, along with that of their managers and co-workers – but not if management retains control of that data and only releases it to them when it supports their own side in such cases.

The top questions for companies looking to install more monitoring algorithms are:

- what purpose is it designed to serve and will that benefit the workforce as well as management?
- is there a non-technological solution to that problem that might be equally as effective?
- are workers made properly aware of this monitoring, consulted about it and giving meaningful consent for it to be introduced?
- who will have access to the data and under what circumstances?
- what privacy and security procedures will be in place?

The second major way in which algorithms can exercise control is by gamification. Such gamification algorithms are particularly prevalent in the gig economy, where workers have at least nominally more control over when and where they work and therefore firms seek to 'nudge' them to doing more work at desired times and places.

One driver for Lyft describes how "a driver's freedom has to be aggressively, if subtly, managed," describing ways in which drivers are

offered regular algorithmically-generated challenges such as "Complete 34 rides between the hours of 5am on Monday and 5am on Sunday to receive a \$63 bonus", with the bonus amount declining the more the driver meets their target, but occasionally offering an unusually lucrative bonus to entice them back if they are absent from work for too long.

These gamified prompts are personally targeted at individual workers based on a whole host of data that the firms gather about them, using AI and analytics to deliver what is calculated to be the most effective incentive to get them to work the hours and places that the company wants them to.

Along with the option to earn meaningless 'badges' for achieving certain targets or unlocking rewards such as discounts, these gamification techniques borrow heavily from the gambling industry's understanding of human addiction psychology. The Lyft driver concludes by saying that "I wanted to be a highly rated driver. And this is the thing that is so brilliant and awful about the gamification of Lyft and Uber: it preys on our desire to be of service, to be liked, to be good. On weeks that I am rated highly, I am more motivated to drive. On weeks that I am rated poorly, I am more motivated to drive. It works on me, even though I know better." (Mason, 2018)

Of course, there may be workers that are happy to work in these gamified environments, seeing the targeted incentive structures as something that helps them achieve higher levels of personal productivity and earnings. Certainly gig economy work is something that many workers do choose voluntarily and while some might find being directed by an algorithm to be dehumanising and taking away their autonomy, others might have the opposite perspective – feeling that the app puts them in control of when and how they work, even if they are strongly incentivised by the algorithms to work at particular times and in particular ways, they still have more freedom than a traditional employee under the direction and control of a line manager.

A study by the Oxford Martin School suggested that the majority over Uber drivers "said they valued flexibility over a salary or fixed hours, and the data showed drivers regularly changed their working hours from week to week (with an average working week of 30 hours)." Despite having below average income for Londoners, the research also found that Uber drivers report higher levels of life satisfaction and worthwhileness than other London workers. (Berger, Frey, Levin and Rao, 2018)

While many of these uses of algorithmic control and incentivisation may be benign or even beneficial, we should still be cautious about the overall direction of travel and where this might lead.

The ultimate expression of monitoring and control by algorithms might have found its expression in China in the form of the social credit system now being rolled out across the country for all citizens. This combines everything from traditional credit scores to facial recognition to track people's movements across the country, to local government records of traffic violations, whether they have correctly sorted their personal recycling or records of cheating in online video games.

Once fully implemented it will be used to exclude people from desirable employment opportunities, ban them from many forms of travel, private schools, hotels, restrict their internet connections and publish their personal data on blacklists. Already as of June 2019 it has been used to refuse 26.8 million air tickets and 6 million high-speed rail tickets to 'untrustworthy' people (Xinhuanet, 2019).

Similarly, looking towards the ability of firms to record and algorithmically analyse ever growing amounts of data about workers, an EU report posits that "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." (European Agency for Safety and Health at Work, 2019)

Reducing what managers use to judge workers to a simple algorithmic score also risks creating a filter bubble around managers, who only end up seeing data the algorithm is collecting, limiting their exposure to a wider range of ideas and eliminating serendipity.

Accuracy and insight

A final important consideration is the degree to which algorithms can increase (or decrease) the accuracy of management judgements

and decisions about the workforce. After all, the goal of many companies adopting algorithmic tools is to improve the quality of their decision making – tools like IBM's Watson computer have demonstrated impressive human-beating capabilities in fields as varied as winning game shows to diagnosing cancer. To do this, however, they need to give serious thought to exactly what tools they are buying and how they intend to use them.

There is some concern about whether many of these workplace algorithms even work at all, at least in the ways they are advertised. Certainly, there are some tools that have proven themselves to be powerful and effective. IBM claim that their patented 'predictive attrition programme' algorithm is now 95% accurate at predicting which workers are about to quit their jobs, saving the company over \$300 million in staff retention costs. (Rosenbaum, 2019)

The way in which using data-driven algorithms completely revolutionised the recruitment approach to the baseball industry in the US after the publication of Moneyball in 2003 is one very niche but high-profile example of the power of a smart algorithmic approach over relying on traditional human intuition. Similarly, technologies such as facial recognition algorithms have improved vastly in the last few years – tests by the US National Institute of Standards and Technology found that the failure rate for identifying a target in a database of 12 million faces dropped from around 5% in 2010 to just 0.1% in 2019 – an error rate low enough to be useful in a workplace context even if it still poses problems for national surveillance use such as in criminal justice.

However, in a rapidly growing and evolving market inflated with hype, and with purchasing managers in companies often lacking in knowledge or understanding of the technology they are buying, there is a real danger of firms investing in, and then trusting in, workplace algorithms that are simply bad at their job. Tech writer and mathematician Cathy O'Neil writes in her book, Weapons of Maths Destruction how a performance management algorithm used in education department of New York City to rate teachers sometimes gave the same teacher a score of 6/100 in one year and 96/100 the next, without any change in their teaching style or anything else to explain the discrepancy. (O'Neil, 2016)

When these measures are being used to determine whether people keep their jobs or not, such discrepancies are a serious cause for concern. Dr Phoebe Moore has cautioned against falling for the hype around some of this technology, saying "there's a whole cadre of marketing experts bloating the discourse in a way that's quite dangerous because it overlooks workers' rights."

Certain technologies in particular have come in for serious criticism for not being scientifically proven. The use of facial scanning algorithms in recruitment, as sold by HireView and in use by companies including Hilton and Unilever, have been described by some experts as little more than 'digital snake oil'. "It's a profoundly disturbing development that we have proprietary technology that claims to differentiate between a productive worker and a worker who isn't fit, based on their facial movements, their tone of voice, their mannerisms," according to Meredith Whittaker, co-founder of the AI Now Institute. "It's pseudoscience." (Harwell, 2019)

Similarly, companies like AC Global Risk are marketing products claiming to determine the level of "risk" posed by people, such as potential employees, based on analysis of their voice during a 10-minute interview that assesses their trustworthiness. Their CEO claims a level of 97% accuracy for these assessments, but independent experts have described their claims as "bogus", saying that "it's currently not possible to tell whether someone is lying... from the voice at greater than 70% accuracy, which is around the same as an average human judgement" and that broader claims about assessing trustworthiness are based on a deeply flawed assumption that concepts like trustworthiness or 'risk' are inherent human traits rather than circumstantial. (Kofman, 2018)

This attempt to assess people's inherent character has been likened to physiognomy, the 19th-century practice of looking for physical signs in the body of moral character or criminality. Similarly, O'Neil concludes by comparing the use of some recruitment algorithms that rely heavily on personality tests as being little better than phrenology – the related 19th-century pseudoscience that purported to assess people's characters' by measuring the shape and bumps on their skulls.

However, we should not get carried away with looking only at the worst performing algorithms. While there certainly are products out there whose performance fails to live up to their marketing hype, this should not tar the reputation of many genuinely impressive tools.

Where algorithms really excel is in processing very large quantities of information that humans would never otherwise be able to do. This also allows algorithmic analysis to spot trends that otherwise couldn't be spotted.

The algorithms from companies like Percolata used to analyse customer footfall patterns in retail stores is one example, correlating them with weather patterns, public holidays, in-store promotions and other factors that would be beyond the ability of human analysts to easily assess. Similarly, performance management algorithms may be able to analyse and discover possible causal factors behind performance and productivity that have eluded us until now. These insights could then be used to fine tune the working environment and improve the quality and productivity of work.

The very best levels of accuracy are often, in fact, achieved by combining humans and AI working together. In an example from the world of healthcare, Harvard researchers found that top AI algorithms can now read cancer diagnostic scans with 92% accuracy, compared with 96% accuracy for top human doctors. However, allowing human doctors to review and work with the AI led to a 99% accuracy rate – higher than either humans or AI working alone.

Similarly, in the world of competitive chess, top chess playing computers have long been able to defeat top human grandmasters, since Deep Blue famously defeated Gary Kasparov back in 1997. However, even today, mixed 'cyborg chess' teams that combine top human players using advanced chess computer tools to assist them, still consistently beat both the best humans and the best computers playing alone. Most interestingly of all, some of the top human-computer chess duos in freestyle tournaments are not even all that highly ranked chess players themselves – the skills needed to most effectively use and work with computer chess software seem to be different from the skills needed to be a great chess player yourself.

The lessons for the workplace seem to be twofold. Firstly, that to get the best quality decisions out of algorithmic management tools, they should be used to advise and work alongside human line managers, not to replace them. Secondly, that to make the best out of these tools, human managers are going to need training in new skillsets, beyond what are typically considered important management skills; learning to understand data and how to make use of it.

Towards more ethical algorithms

For better or for worse (or, more likely, both), AI and algorithms in the workplace are only going to become more prevalent in the years ahead. As this happens, we must make sure not to neglect the importance of taking an ethical approach to their deployment. Elevating a system of automated decision making above humans on the basis that it is somehow 'smarter' than human decision makers is a dangerous road to set out on. Once we start assuming that machines are bound to make objectively better decisions than us without bias or error, there is no room left for a discussion of ethics.

UNI Global Union, an international workforce organisation representing 20 million global workers, has published what they consider to be a list of 'top 10 principles for ethical artificial intelligence' to ensure algorithms have an ethical impact on people and society (UNI Global Union, 2017). Several of these are of particular interest to the debate around algorithms in a workplace context:

- demanding transparency
- equipping AI systems with an 'ethical black box' that records their decisions and reasons
- adopting a human-in-command approach
- banning the attribution of responsibility to robots
- sharing the benefits of AI systems
- ensuring a genderless, unbiased AI

To this we would also add one other vital principle – ask why you need the technology, or whether you really need it at all. Algorithms

have potential to support and improve the quality of a whole host of workplace decisions, but companies should always ask themselves what they are actually going to get from an algorithm that they couldn't get through genuine human discussions with their workers instead. If the answer is 'not much' other than that it seems easier to buy an algorithm, while having genuine human discussions appears quite hard, then perhaps algorithmic management is not the right approach to take.

Humans in charge - transparency, accountability and redress

If workers are to have any confidence that algorithms are being used in an ethical way, it's vital that there is openness about what they are being used for and how they come to their decisions. Many companies have to date been far too secretive about what their algorithms are doing, often on grounds that the source code is proprietary information, and sometimes leading to the not unreasonable suspicion that managers themselves might not even really know how the tools they are using work.

Transparency, however, is not simply about sharing the source code. In fact, in many cases sharing the code would not be helpful – it is all too easy to conceal what is actually happening beneath a level of obfuscating complexity. Rather what is needed is simple, understandable explanations in ordinary English that workers, consumers and others can follow as to why certain decisions have been made. If one worker is allocated a pay rise or bonus by algorithm and another is not, both workers have a right to know why in terms that make sense to them, rather than simply be told "the algorithm said so".

UNI Global Union recommends that all workplace algorithms be fitted with an ethical 'black box' that would provide an explainable record of all its decisions. "Applied to robots, the ethical black box would record all decisions, its bases for decision making, movements, and sensory data for its robot host. The data provided by the black box could also assist robots in explaining their actions in language human users can understand, fostering better relationships and improving the user experience." (UNI Global Union, 2017)

Greater transparency will also help in turn to drive faster improvements in accuracy and in bias reduction. Where workers and managers can see that algorithms are leading to biased or unfair decisions, this will allow them to not only correct the individual decision but also to examine ways in which the algorithm can be improved to reduce the chance of similar mistakes in the future.

There is also a need for not only internal but external accountability. Algorithms making decisions that have significant impacts on workers' lives should be regularly and externally tested and assessed by independent third parties – for both accuracy and for bias. This can be done by trusted organisations that maintain the confidentiality of any genuinely proprietary code or trade secrets.

Where transparency does have a limit is where it comes up against personal privacy. Not the right of a company to keep its decision-making process private, but the right of individual workers to privacy and the right to a personal life. There is no good reason for algorithms to collect data on workers that might extend beyond their legitimate workplace activities and into the private sphere of their conversations with colleagues, their personal lifestyle or their activities at home.

As for the human in command principle, this may be the most important factor of all. Where algorithms make mistakes – and they will – it is absolutely essential that there is a human with oversight of the process who can step in, correct the mistakes and provide redress to anyone affected. Ultimately a human manager with the ability to both overrule individual decisions and request changes to the source code should be held responsible for the outcome of all algorithm-supported decisions. A machine is incapable of being morally or legally responsible for decisions and companies should avoid ever attributing the responsibility for decisions to non-human entities.

Involving the workforce

An ethical approach to algorithms at work must also be one that involves the workforce, both in consultation about their introduction and in partnership over their use and evaluation. At present, surveys suggest that there are fears from trade unions and many parts of the workforce that algorithms are something being imposed on workers rather than adopted in collaboration with them.

The best way that this fear can be addressed is by making sure that algorithmic tools are being adopted in a way that benefits both sides. If algorithms are being brought in to achieve productivity gains, make sure that those gains are shared with the workforce in terms either of more pay, or shorter working hours. If algorithms are being brought in to make HR processes more convenient, make sure they improve convenience for both frontline workers and HR managers. If companies are bringing in performance review algorithms, make sure they are used to provide helpful feedback to workers and target support at those who need it, rather than just used to put more pressure on workers to work harder, longer and faster.

Where algorithms are used to monitor and gather data on workplace activities, make sure workers (subject to privacy and security concerns) can access that data in a way that is useful to them, not just reserving valuable insights and records for senior managers.

As with all technological change, the change management process will not succeed if it is simply a top-down approach that doesn't secure buy-in from the workforce. While many of the tools discussed above, like shift scheduling algorithms, do have the potential to benefit both workforce and management, that doesn't mean that workers won't be sceptical or hostile towards such technology if it's imposed on them without their consent. Securing this buy-in requires a clear explanation of the reason why the tools are needed, what they are going to achieve and how they are going to operate, as well as what changes this will bring for workers' day-to-day lives.

Workers, particularly younger workers, are highly optimistic about the future of technology and keen to embrace it. A survey of Prospect members in 2018 found that 54% of those aged 25 and under were optimistic about technology, compared with only 5% who were pessimistic (Prospect, 2018).

However, this general optimism will not automatically translate into support for individual algorithmic tools that managers seek to introduce, unless they have taken the time to lay the groundwork of support building with the workforce from an early stage. Managers should not simply assume that if workers have questions or concerns about the introduction of algorithmic management tools that it is because the workforce is full of luddites or otherwise resistant to new technology in principle. Taking the time to address what specific concerns they have will pay dividends in terms of a smoother rollout of the new systems.

Who controls the data?

For AI and algorithms at work to achieve a better workplace it is essential that access and ownership of data is properly safeguarded. Control of data is set to be one of the great social and economic battlegrounds of the 21st-century global economy. Access to and control of data is probably the single most valuable resource in the modern world.

Estonia, one of the most digitalised countries in the world with a very advanced and extensive e-government system for all citizens, also has one of the most developed set of data control safeguards in the world. All citizens are able to decide what of their data is available to whom and how it can be used. In order to ensure high ethical standards, the same principle should extend to workers' access to and control of their data in a workplace context.

As well as their general principles of ethical approaches to AI, UNI Global Union have set out 10 main principles for workers' data protection (UNI Global Union, 2017):

1. "Workers and their union representatives must have the right to access, influence, edit and delete data that is collected on them and via their work processes."
2. Data processing safeguards must ensure that workers are properly informed and consulted before being monitored or having

their data processed.

3. The data minimization principle: "Collect data and only the right data for the right purposes and only the right purposes, to be used by the right people and only the right people and for the appropriate amount of time and only the appropriate amount of time."
4. Transparency – workers need to know details about what data is being processed, when, where, why, how and by whom.
5. Respect for relevant laws and human rights, including the ILO and UN's statements of rights
6. The right to an explanation – of why or on what basis any decisions affecting them were made.
7. The exclusion of biometric or other personally identifiable information from any data processing unless absolutely necessary and done in accordance with strict scientific methods and security principles
8. The use of equipment revealing employees' locations is particularly sensitive and should be restricted to cases where it proves absolutely necessary to achieve a different legitimate purpose.
9. The establishment of a multidisciplinary, intercompany data governance body to oversee all processing of worker data.
10. A collective agreement with the workforce, concluded at company or sectoral level through collective bargaining, to cover all of the above points.

Given the importance of personal data to people's overall wellbeing at work, and their sense of personal autonomy, none of these principles should be considered too radical or demanding for businesses to accept. Any firm with the resources to invest in algorithmic management tools also has the resources needed to make sure that this is done in an ethical manner.

Overall, as time passes and our understanding of algorithms improves and errors are corrected, the areas where it is safe, responsible and beneficial to deploy algorithms at work should expand. While there may be some simple and routine parts of HR and people management where it makes sense to replace humans altogether, there will always be benefits to having human line managers to supervise, support and develop workers – not only because it is less alienating than being managed by an algorithm, but also because there are genuine human skills of empathy and communication that are vital to good management and which machines cannot replicate.

In many cases, though, the very best decisions will be those taken by fusing the best parts of both human and machine – signed off and ultimately accountable to human managers, but supported and advised by AI and other algorithmic tools. Overall our attitude to the development of these tools should be one of optimism tempered by caution. The role of ethics is absolutely vital and we should never rush to embrace technology for its own sake without taking due account of the dangers of bias, fairness, autonomy or other ethical issues. In the end it's important to remember that algorithms, as with people, are unlikely ever to be perfect at what they do but with the right care and attention they can be a force for good.

Bibliography

Access Now (2019). 'One year under the EU GDPR: an implementation progress report' Access Now.

Alsever J (2016, March 21). 'Is software better at managing people than you are?' Fortune.

Barocas S and Selbst AD (2016). 'Big data's disparate impact' California Law Review.

Berger T, Frey CB, Levin G and Rao S (2018). '[Uber happy? work and wellbeing in the "gig economy"](#)' Working paper to be presented at the 68th Panel Meeting of Economic Policy in October 2018, Oxford Martin School, University of Oxford.

Bertrand M and Mullainathan S (2004). 'Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination' American Economic Review.

Bogen M and Rieke A (2018). ['Help wanted: an examination of Hiring Algorithms, Equity, and Bias'](#) Upturn.

Booth R (2019, April 7). ['UK businesses using artificial intelligence to monitor staff activity'](#) The Guardian.

Bri ne P (2017). ['Mind over Machines: New technology and employment relations'](#) Acas.

Brynjolfsson E (2018, May 18). ['Where Humans Meet Machines: Intuition, Expertise and Learning'](#) Medium.com. (D. Kahneman, Interviewer)

Buranyi S (2018, March 4). ['How to persuade a robot that you should get the job'](#) retrieved from The Observer.

Bush S (2019, January 25). 'Of course algorithms are racist. They're made by people' New Statesman.

Collins L, Fineman D and Tsuchida A (2017). ['People analytics: Recalculating the route.'](#) Global Human Capital Trends, Deloitte Insights.

Courtland R (2018, June 20). ['Bias detectives: the researchers striving to make algorithms fair'](#). Nature.

Dastin J (2018, October 10). ['Amazon scraps secret AI recruiting tool that showed bias against women'](#) Reuters.

Dellot B (2017, November 10). ['The Algorithmic Workplace Need Not Be A Dystopia'](#) RSA Blog.

Eder S (2018, July 31). ['Should you use AI for performance review'](#) LinkedIn.

European Agency for Safety and Health at Work. (2019). OSH and the future of work: Benefits and risks of artificial intelligence tools in workplaces.

Fisher A (2019, July 14). An Algorithm May Decide Your Next Pay Raise. Fortune.

Fry H (2018). Hello World: How to Be Human in the Age of the Machine. Transworld Publishers.

Google. (2017). ['People Analytics.'](#) Retrieved from ReWork.

Harwell D (2019, November 6). A face-scanning algorithm increasingly decides whether you deserve the job. Washington Post.

Heric M (2018). HR's New Digital Mandate. Bain & Company.

Hopping C (2015, June 5). The truth about talent selection algorithms. Retrieved from Launch Pad Recruits.

Hume K, and LaPlante A (2019, October 30). 'Managing bias and risk at every step of the AI-building process' Harvard Business Review.

Information Commissioner's Office. (2019, May 22). ['Rights related to automated decision making including profiling'](#) retrieved from ICO.

- Kobie N (2017, January 16). Workplace monitoring: would you let your boss track your mood. IT Pro.
- Kofman A (2018, November 25). ['The Dangerous Junk Science of Vocal Risk Assessment'](#). The Intercept.
- Lambrecht A and Tucker C (2018). 'Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Adverts'.
- Lee MK, Kusbit D, Metsky E and Dabbish L (2015). 'Working with Machines: The Impact of Algorithmic and Data-Driven Management on Human Workers.' Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.
- Leicht-Deobald, U, Busch T, Schank C, Weibel A, Schafheitle S, Wildhaber I and Kasper G (2019). 'The Challenges of Algorithm-Based HR Decision-Making for Personal Integrity.' Journal of Business Ethics.
- Levy KE (2015). 'The Contexts of Control: Information, Power and Truck-Driving Work.' The Information Society(31:2).
- Marvin R (2019). ['The Best Employee Monitoring Software 2019'](#) retrieved from PC Magazine.
- Mason S (2018, November 20). 'High score, low pay: why the gig economy loves gamification'. The Guardian.
- Mateescu, A and Nguyen, A (2019). ['Algorithmic Management in the Workplace'](#) Data & Society.
- Miller, AP (2018, July 26). 'Want Less Biased Decisions? Use Algorithms' Harvard Business Review.
- Moore, P V (2018). 'The Threat of Physical and Psychosocial Violence and Harassment in Digitalized Work' International Labour Organization.
- Moore, PV and Joyce, S (2018). 'Black box or hidden abode? Control and resistance in digitalized management' Lausanne University workshop 'Digitalization and the Reconfiguration of Labour Governance in the Global Economy'.
- O'Connor, S (2016, September 8). ['When your boss is an algorithm'](#) Financial Times.
- O'Neil, C (2016). 'Weapons of Math Destruction' Crown Books.
- Parise S, Kiesler S, Sproull L and Waters K (1999). 'Co-operating with life-like interface agents' Computers in Human Behaviour.
- Prospect. (2018). Prospect survey of members, 2018.
- Prospect. (2019, September). Written evidence submitted by Prospect Union (AFW0047) to [BEIS Select Committee Report on Automation and the Future of Work](#).
- Rapp N and O'Keefe B (2018, January 8). ['These 100 companies are leading the way in AI'](#) Fortune.
- Roose K (2019, June 23). ['A machine may not take your Job, but one could become Your Boss'](#) New York Times.
- Rosenbaum E (2019, April 3). ['IBM artificial intelligence can predict with 95% accuracy which workers are about to quit their jobs'](#) CNBC.
- Rosenblat A, Levy K, Barocas S and Hwang, T (2016). 'Discriminating Tastes: Customer Ratings as Vehicles for Bias' Data & Society.

The Economist. (2018, June 21). How an algorithm may decide your career.

TribePad. (2019). 'Hiring humans vs recruitment robots: How technology is changing recruitment and what that means for your job'.

TUC. (2018). 'I'll be watching you; A report on workplace monitoring'.

UNI Global Union. (2017). ['Top 10 Principles for Ethical Artificial Intelligence'](#).

UNI Global Union. (2017). ['Top 10 Principles for workers' data privacy and protection'](#).

University of Hertfordshire. (2019). ['Platform Work in the UK 2016-2019'](#). Statistical Services and Consultancy Unit, University of Hertfordshire and Business School.

Walsh M (2019, May 8). 'When algorithms make managers worse' Harvard Business Review.

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

Xinhuanet. (2019, July 17). 2682?????????????. Retrieved from xinhuanet.com China.