Future of Work

Investors' Expectations of Ethical Artificial Intelligence in Human Capital Management

HSBC Asset Management | MACFARLANES | LSE The Inclusion Initiative

Foreword

"To understand the causes of things, for the betterment of society."

It is in this spirit that the three authors embarked on their journeys to write this paper, first conceptualised in August 2021.

The paper is intended for investor engagement with companies on the responsible use of AI in hiring and workforce management.

At the outset, this paper sought diverse and inclusive viewpoints from a range of stakeholders of the responsible investment ecosystem, for the paper is intended to provide investors' expectations on ethical artificial intelligence (AI) applications in human capital management.

As such, across the paper, you will find expert opinion, insights based on independent surveys and the voices of diverse employees.

We hope you enjoy reading this paper and find practical tips on applying human-centric AI. Ultimately, we invite you to join us in making workplace more transparent and inclusive.

Thank you.

From the authors

Why should I read this paper?

From a Regulatory Perspective:

"Carefully designed and properly used, I believe that AI can advance diversity and inclusion in the workplace by mitigating the risk of unlawful discrimination. At the same time, poorly designed and carelessly implemented, AI can discriminate on a scale and magnitude far greater than any individual HR professional. As a public servant, I am committed to ensuring that AI helps eliminate rather than exacerbate discrimination in the workplace and provides the clarity that workers, employers, data scientists, and policymakers have long been asking. I applaud efforts, like this one, that help educate our stakeholders in identifying ways to accomplish our goals of preventing and remedying unlawful employment discrimination and advancing equal opportunity for all in the workplace."

Commissioner Keith Sonderling, US Equal Employment Opportunity Commission

From a Health and Well Being Perspective:

"The deployment of AI systems has outpaced research into the mental health dimension of their widespread use. Relevant issues include the implications to mental wellbeing from privacy violations, as can happen from workforce surveillance, given privacy's role in mediating crucial psychological functions like resilience and the ability to bounce back from adversity. Further, AI-mediated psychotherapy is a nascent field within telemedicine that may over time--but can't quite yet--offer guidance on the ability to analyze facial expression and assess personality in non-clinical settings. This paper is a comprehensive, psychologically-informed take on human capital management, one that doesn't avoid asking the difficult questions that must be asked for the field to evolve ethically and in a data-driven way."

Professor Elias Aboujaoude, MD, MA, Stanford University Department of Psychiatry and Stanford University Human-Centered AI

From a Civil Society Perspective:

"This timely, informative paper makes a powerful case in support of a diligent approach to the use of artificial intelligence in companies' human resource management. By adopting an approach to AI that is grounded in clear ethical principles, companies will not merely be doing the right thing, they will protect themselves and their investors from regulatory, reputational and financial risks.

The issues set out in this paper are complex, and the uses of AI in employment practice are rapidly evolving, penetrating deeper into companies' own operations and critical supply chains in a growing range of sectors. The pandemic has accelerated the trend, heightening both the risks and the opportunities for employers of adopting AI. In this context, the paper offers a coherent framework for tackling the emerging challenges and provides investors with key questions to pose to companies, not least in relation to AI's impacts on diversity, equity and inclusion. An obvious starting point for investors wanting comfort that risks are being appropriately managed is to ask about the extent of C-suite engagement and board oversight of AI as it pertains to employment and wider workforce practice.

It is the very essence of high-quality responsible investment practice to recognise and get to grips with the implications of complex ESG issues earlier than other actors in capital markets. HSBC Asset Management and the London School of Economics are to be commended for engaging fully with this fascinating and ethically demanding terrain, and to share their insights with others in the market by publishing this paper. Institutional investors acting on behalf of working people who are saving for their retirement should pay close attention to the topic given its relevance both to those people's lives at work and to the protection of their financial assets from poorly managed risks. I would hope to see ethical AI in human capital management emerge as a key focus area for responsible stewards of financial capital. This paper lays the ground for that development and should be widely read by corporate leaders and professional investors."

Catherine Howarth, CEO of Share Action

From a Stock Exchange Perspective:

The Future of Work: Investors Expectations on Ethical AI in Human Capital Management is a candid review of the current state of an AI field that is deemed to have significant future potential but is fraught with risk if incorrectly employed. The emergence of artificial intelligence possibilities in business has been brought about by three major factors that have occurred over the past two decades. The three factors are the implementation of advanced hardware, the ability to manage large amounts of data (sometimes referred to as big data) and the emergence of new algorithms for the manipulation of the data. While these new technologies offer the possibility of tremendous benefits in the future what has not changed is human beings. Our people are complex beings and their behaviour is often beyond the understanding of AI in its current form. Where AI does work well is as a tool when decision rules are quantifiable and computers can be used to interact almost instantaneously with vast amounts of data. At the stock exchange we used sophisticated algorithms as a tool to assist our staff to identify stock trades that might contravene laws, regulations or our exchange policies. The growth of massive numbers of messages(orders) and the speed (micro seconds) which they arrived made it impossible for human beings to adequately ensure trading was fair. Only computers equipped with AI could identify the improper trade within the overall trading environment.

However, when it comes to management of people (our human capital) applications are at the earliest stage of development. As the paper discusses in several sections the quality of the AI application is highly dependent upon its programmer's input. Applications purchased off the shelf which are essentially black boxes of programming by outsiders should never be trusted by the organization until thoroughly tested, and the organization is confident that they know exactly how the application functions. Human beings are complex individuals and interactions with them cannot be programmed in the same way as the trading process at an exchange. Despite these concerns the future will see more advances in the use of AI for working with people and we should all strive to ensure this is accomplished in a way that improves outcomes on a fair and inclusive basis.

Richard Nesbitt, Professor, Rotman School of Management and Chair of the Inclusion Initiative at LSE. He is former CEO of the Toronto Stock Exchange and COO of CIBC.

From a Corporate Disclosure Perspective:

"Artificial intelligence and human capital management have historically been complete separate domains in terms of issues and expert. But they are converging rapidly as professionals charged with human capital management are increasingly relying on AI tools in making their decisions. Like all tools, AI can be used for good or ill and often a complex combination of the two. Getting more of the former and less of the latter means good governance around AI. Good governance, in turn, has long been the domain of investors in their engagements with companies. This paper provides a solid foundation to help investors do this in an effective way."

Professor Robert G. Eccles, Saïd Business School, University of Oxford. Founding Chairman of the Sustainability Accounting Standards Board (SASB) and one of the founders of the International Integrated Reporting Council (IIRC)

From a Sustainable Development and Corporate Benchmarking perspective:

"Digital technologies such as AI can either help or hinder the achievement of the 17 UN Sustainable Development Goals. While AI has much potential to augment human capacities to solve global challenges, it can also derail progress, especially when applied without care in areas such as human capital management. In particular, the achievement of SDG 5 (Gender Equality) and SDG 10 (Reduced Inequalities) are endangered by unprincipled and unthinking applications of AI towards workforce acquisition, management, and optimisation. This insight-packed paper is therefore a really important contribution guiding investors on these issues. Investors have a key role in facilitating the achievement of the SDGs and ensuring that digital transformation happens in an inclusive and trustworthy way. The 2021 findings of the World Benchmarking Alliance's Digital Inclusion Benchmark unfortunately show that awareness and recognition of the risks among digital technology companies continue to be low, with only 20 out of 150 companies demonstrating any public and general commitment to ethical AI. This paper gives hope that many investors will be willing to use their influence for positive change."

Lourdes O. Montenegro, Digital Sector Transformation Lead, World Benchmarking Alliance

From a Responsible Investor's Perspective:

"This is an excellent and thought-provoking paper. An insightful mix of critical reflection, case study and practical recommendation. It establishes a clear foundation for companies to use their leverage to create 'gold standards' for AI systems that seek to enable the flourishing of all humans – both individually and collectively."

Anna McDonald, Secretary, The Church of England Ethical Investment Advisory Group

"Digital transformation is critical to the long term success of any company. When we use technology, especially machine learning technologies, to empower people, it should be used to improve fairness, efficiency, transparency and accountability. At HSBC Asset Management, our diversity, equity and inclusion (DE&I) programme has twelve work streams, providing our colleagues a connected platform to take part in driving DE&I goals. This paper sets out our expectations for investee companies in human capital management, a rapidly developing area of machine learning application. We want to ensure that companies not only focus on establishing relevant digital ethics principles, which is important, but also have clear plans for implementation in important areas where fairness, efficiency, transparency and accountability can be improved. With increased focus on DE&I, this is an area that investors should drive the use of concrete performance indicators. This paper provides useful suggestions on key performance indicators for companies that are planning to, or are using, advanced technology and machine learning systems in human capital management practices."

Stuart White, Chief Executive Officer UK & International; and Global Head of Strategy, HSBC Asset Management

"As the world embraces the Fourth Industrial Revolution, otherwise known as the digital revolution, the governance of ethical AI in human capital management is a major investor priority in their engagements with companies. At the same time, the world's movement towards a 'just' transition to combat climate change, coupled with the COVID crisis, has elevated the importance of ordinary workers as pivotal to a company's long-term success, particularly as our economies have shifted away from manufacturing to knowledge-based industries. Corporate boards must be cognisant of these issues and the risks and opportunities of ethical AI in human capital management – ultimately impacting a company's efficiency, productivity, innovation, and competitiveness. In addressing this, the Future of Work paper artfully describes optimal governance parameters, regulatory drivers and key determinants for higher standards and corporate practices worldwide."

Kerrie Waring, CEO, International Corporate Governance Network (ICGN)

Acknowledgements of interviewees and contributors

- Commissioner Keith Sonderling, Rachel See and Tara Helfman, the U.S. Equal Employment Opportunity Commission (EEOC)
- Anna McDonald, Secretary, The Church of England Ethical Investment Advisory Group
- Dr Michael Natusch, Chief Science Officer, Prudential Plc
- Professor Elias Aboujaoude, MD, MA of Stanford University and its Human-centered AI (HAI) initiative
- Dr Juergen Rahmel, Head of Digital Technology and Innovation, Compliance APAC, HSBC Group
- Dr Mark McDonald, Head of Data Science and Analytics, HSBC Global Markets and Research
- Nicholas Dowell, portfolio manager and Rory Muldowney, investment analyst from Neurodiverse Ability workstream of the Diversity, Equity and Inclusion (DEI) initiative, HSBC Asset Management
- Simon Rawson, Aine Clarke and Charlotte Lush, ShareAction and its Workforce Disclosure Initiative (WDI)
- Dr Frida Polli, CEO, Pymetrics
- Dr Anshul Sheopuri, Vice President, Chief Technology Officer, Data & AI, IBM
- Patti Nichols, Data & AI, HR Ethical AI Lead, IBM Human Resources
- Nickle LaMoreaux, Senior Vice President & Chief Human Resources Officer, IBM
- Dean Frankle, Managing Director and Partner, Boston Consulting Group (BCG)
- Sarah Maynard and Rebecca Fender, The CFA Institute
- Diane Gherson, Harvard Business School, Board Member of National Academy of Human Resources and former Chief Human Resources Officer, IBM

Table of contents

Beginning with a Little Story …

> "AI is neither artificial or intelligent"
>
> **- Kate Crawford, Principal Researcher, Microsoft Research**

During COVID-19 lockdown, a friend called.

"Do you think I was speaking to an AI in my values alignment assessment interview?" she asked.

"What makes you think that?" I said, wondering if another algorithm has allegedly 'passed' the Turing test - a test of a machine's ability to exhibit intelligent behaviour equivalent to, or indistinguishable from, that of a human. After all, the language model GPT-3[1] , caused a lot of excitement when it was unveiled[2] in 2020 by Open AI[3].

"The interviewer was so good at reciting what I told her. Basically, she asked me for an example that demonstrates a specific value, such as integrity. I gave a situational example, then she summarised what I said, perfectly. After which, I acknowledged her summary."

Her face dropped – I could see that because we were on a Zoom call.

"Did I just spend an hour talking to a machine? I did not see her face. There was no video, just a voice. She was polite, in a clinical way."

"It is technically possible that they use a transcription service to record your answer, and an AI language model summarised what you said. The Open AI team that created GPT-3 did say that it is good at summarising text, especially with the user giving feedback[4].. In any case, if you get the job then you will find out, won't you?" I replied, hoping that she gets the job.

[1] The full name of GPT-3 is Generative Pre-trained Transformer 3
[2] https://arxiv.org/abs/2005.14165
[3] https://openai.com/blog/openai-api/
[4] https://openai.com/blog/learning-to-summarize-with-human-feedback/ and Hutson M (2021) Robo-writers: the rise and risks of language-generating AI Nature 591, 22-25. See: https://www.nature.com/articles/d41586-021-00530-0

Key Recommendations

Artificial Intelligence is tremendously useful, when applied correctly, to improve fairness, transparency, efficiency and even to correct historical biases, including ethnic and gender biases imposed by humans. However, when applied in a 'plug and play' manner, we risk exacerbating social and economic inequality.

We recommend that:

- Companies should create a machine learning platform to organise data, collaborate amongst data scientists and diverse assessment team members with domain expertise to evaluate models and enable deployment. Where possible, companies should create automated and repeatable data preparation processes and feature engineering to ensure consistency.
- The platform should have a data versioning function to keep track of changes in data used over time, different model experiment and test results. There should be a feedback loop and associated documentation that integrate outcomes and lessons learnt to improve performance over time.
- When engaging with companies, investors can ask for the disclosure and explanation of the following key performance indicators (KPI): to attest that they are employing ethical - fair, inclusive, compliant and trusted - artificial intelligence technologies that promote diversity, equity and inclusion (DEI) in the workplace.
- Companies should include a means of redress and clear remediation mechanisms to an accountable human.

Hiring	Key Performance Indicators (KPIs)
Workforce planning	• Explainability: The company can explain how often AI models are updated and back-tested for relevance. • The company keeps a library of the customised AI models used (if any) and document the process of customisation and changes over time. • Transparency: The company documents how the role models of specific jobs and skills are selected for customised training (if any) of vendor provided AI.
Recruitment	• Research that forms the foundation of the hiring models readily accessible by employer (clients) and candidates (subjects/users). • Research should be peer-reviewed and research data independently verified where possible. • Company can articulate the source of the training data and explain the relevance of its use to the company specific use cases, and highlight disparity. • Evidence that vendors incorporate data privacy protection by design principles into their AI systems. • The company can explain how the AI system predicts and maximises an outcome sought as a purpose of the recruitment exercise. • **See case study 1: Boston Consulting Group.**
On-boarding	• Documented benefits of the automation of on-boarding tasks. • Satisfactory surveys of AI-enabled on-boarding experience. • A feedback loop that integrates survey comments to improve employee experience. • **See case study 2: Citizens Bank**.

Culture	Key Performance Indicators (KPIs)
Competency management	• Company demonstrates that it is more effective when using AI to catalogue, manage and develop employees' skills, which leads to beneficial outcomes such as more targeted learning, improved mobility, more transparency and improved employee satisfaction with the performance assessment processes.
Team collaboration	• There is a process to ensure consent from employees is obtained, and the overarching principle is to opt-in and not opt-out. • Evidence that human rights and ethical impact assessments have been conducted at vendors. • Companies can articulate how it assesses the challenge of 'gaming the influencer system'. • **See case study 3: Nestle.**
Mobility	• Measure attrition rates of mobile staff versus non-mobile staff. • Measure satisfaction rates of mobile staff versus non-mobile staff. • Measure career progression of mobile staff versus non-mobile staff.
Performance	• **Key Performance Indicators (KPIs)**
Remuneration	• Employees understand how their performance is evaluated including the factors used and relative weightings of those factors. • Satisfaction surveys show that employees believe an AI enabled system brings more transparency and fairness in remuneration.
Attrition	• Companies can articulate what are the key metrics and outcomes used to evaluate performance, promotion and productivity. • Attrition rates reduce as a result of better prediction of performance and timely actions to at-risk valued employees who may leave.
Learning	• Personalised learning programmes are skills based, role based and both to ensure flexibility and transferability of skills are supported e.g. no role-based stereotypes as constraints in learning modules. • Company demonstrates that learning recommendations are integrated with mobility performance and remuneration. • **See case study 4: IBM.**

Executive Summary

COVID-19 pandemic has accelerated the digital transformation of businesses. This transformation is not only changing customers' behaviours, supply chain and logistics arrangements; it is also changing hiring practices, communication norms, team dynamics and performance evaluation processes.

This paper consists of five key chapters to explore the different aspects of how ethical artificial intelligence (AI) principles and practices can be integrated into human capital management: (1) Governance; (2) Legal and Regulatory Considerations; (3) Hiring; (4) Culture; and (5) Performance.

Chapter 1 Governance highlights that although the adoption of AI applications is often based on bottom-up assessment and business needs, the rapid development of technologies requires a comprehensive AI embedded strategy and risk management process, led by the C-level executives with board oversight, especially when such technologies are involved in tasks that have significant human impact. We make four key recommendations on how to strengthen governance. We highlight that correcting for data bias ex post is possible. From a human-centred perspective, whilst choosing an opt-in led approach ensures consent, it might also reduce appropriate representation in the data.

Chapter 2 Legal and Regulatory Considerations , prepared by Mark Lewis of Macfarlanes LLP, outlines the implications of failure to comply with law and regulation, board accountability and governance considerations, regulatory interest, concerns and initiatives in this area, and how and why regulatory and legal challenges may arise in the context of AI in human capital management. It goes on to provide representative examples of: (1) current laws and regulations that apply specifically and directly to the use of AI in the workplace, (2) prospective law and regulation that is expected to apply specifically and directly in this context (focusing on the European Union's proposed Artificial Intelligence Act), and (3) current laws and regulations from selected jurisdictions that we expect to apply to the deployment of AI in the workplace by judicial interpretation and application, this last category arranged thematically and aiming to show the interplay between those laws and regulations and AI systems, processes and, where applicable, data sets. The themes in (3) are:

- the fundamental legal basis of the employer-employee relationship,
- employee rights not to be unfairly dismissed,
- equality laws, and discrimination in hiring processes, based on certain characteristics,
- privacy rights and laws governing surveillance in the workplace,
- data protection,
- rights to associate and to collective bargaining and representation, and
- health and safety at work.

The **Findings on Workforce Surveillance** section, **prepared by Charlotte Lush of ShareAction,** suggests that companies have already been gathering some data on the use of workforce surveillance internally, but this data is yet to be incorporated in well-established sustainability reporting requirements. In general, few companies are involving workers in their surveillance measures, although companies that conducted more surveillance sought more worker engagement and input in data collection processes. For companies to be able to genuinely respect workers right to privacy, worker involvement is crucial.

Chapter 3 Hiring is divided into three sections – workforce planning, recruitment and on-boarding. Workforce planning is a core business process which aligns changing organisation needs with people strategy. This involves identifying skills gaps, future skills need and disruptive intervention necessary to take a company forward, Using **Boston Consulting Group** as a case study, we reflect on the opportunities and challenges of workforce planning from the perspective of accountability, grounding from literature and independent AI audits and data bias risks. From a human-centred perspective, fairness has many different interpretations. The most commonly accepted form of fairness definition is 'disparate impact'– refers to practices in employment, housing, and other areas that adversely affect one group of people of a protected characteristic more than another, even though rules applied by employers or landlords are formally neutral.

Recruitment is a process that covers role creation, writing of job description, role advertising, resume screening, shortlisting, interviews (including aptitude tests, language skills test and case studies, where appropriate), psychometric and cognitive tests and corporate values alignment assessment. We cover AI in candidate search, resume screening, cognitive tests and video interviews, highlighting the opportunities and risks in each approach using case studies.

On-boarding is an essential aspect of employee recruitment. Successful on-boarding means enabling the new colleague to 'hit the ground running', generating good first impression of the employer, and accelerating cultural integration. We use **Citizens Bank** Jamie™ Conversational AI Virtual Assistant as a case study.

Chapter 4 Culture is divided into three sections – competency management, team collaboration and mobility. Competency management refers to systemically cataloguing, managing, and developing of job related skills. Soft and hard skills mapping encourage mobility between different types of jobs, which also encourages agility amongst workforce and allow cross learning between departments, strengthening culture and collaboration across different businesses.

The use of employee network analytics has increased during COVID-19 when team interactions moved from in-person to online experience where interactions can be tracked and monitored. Using **Nestle** as a case study, we highlight that workplace analytics can deliver impromptu and timely communication. However, social influencers campaigns may put unexpected pressure on colleagues to be more 'sociable' than their authentic self. From a human-centred perspective, workplace analytics may also unintentionally benefit extroverts over introverts; technology savvy employees over those who are less technology-oriented.

Chapter 5 Performance is divided into three sections – remuneration, attrition and learning. Currently, most pay and promotion decisions are made by managers who work closely with the individuals they manage. However, unconscious biases can creep into these decisions when the impact of likeability or similarity can hinder objectivity in these decisions. When a multitude of factors influence the pay of an employee, algorithms can be a tool to translate them into quantified outcomes. Furthermore, algorithms can take into account the supply & demand of certain employee skills based on labour market trends and compensate employees accordingly. Positively, AI can improve employee perceptions of fairness in performance evaluation through transparency and remove pay and promotion gaps; however, from a human-centred perspective, there are risks that some factors contributing to performance may not be quantifiable and therefore cannot be 'programmed' accurately.

AI in learning is often a 'self-help' function for individual employees. The most common technique used is user item recommendation, a collaborative filtering technique to identify similar users based on items that are common. Using **IBM** as a case study, we highlight how AI – defined as advanced classification systems that optimise outcomes based on user-defined factors - can guide our decisions on human capital management.

Introduction

COVID-19 pandemic has accelerated the digital transformation of businesses. This transformation is not only changing customers' behaviours, supply chain and logistics arrangements; it is also changing hiring practices, communication norms, team dynamics and performance evaluation processes. For example, LinkedIn surveys[5] suggest that 81% of talent professionals believe virtual recruitment is here to stay post COVID. This is aligned with an established trend of increased use in people analytics[6].

Digital team chats, shaping culture, generating workplace interaction and activity data. More webinars and meetings are recorded and saved on shared drives and on social media. A few will be watched again, but artificial intelligence ("AI") algorithms can analyse these recordings; AI is tasked with evaluating and scoring behaviours without ever having any meaningful interactions with us, as a human being. What are the implications on human capital management in this aspect of digital transformation?

The impetus to author this paper comes from a UK All Party Parliamentary Group ("APPG") on AI discussion and recommendations to House of Lords and House of Commons members on the UK National AI Strategy[7].

"When looking at thematic applications, such as meeting diversity, equity and inclusion (DEI) goals, do we spend enough time thinking about data bias? The AI value chain is very long. Have we put enough effort into understanding the quality of data, the labelling of data, the classification of data that affects the standards of algorithmic training, which in turn impacts performance of our analytics?"

The APPG for the Future of Work[8] Inquiry into AI at Work organised three enquiries and published The New Frontier: Artificial Intelligence at Work[9] in November 2021. The report finds that pervasive monitoring and target setting technologies, in particular, are associated with pronounced negative impacts on mental and physical wellbeing as workers' experience the extreme pressure of constant, real-time micro-management and automated assessment. A core source of anxiety is a pronounced sense of unfairness and lack of agency around automated decisions that determine access or fundamental aspects of work. The challenges identified lie between data protection, labour and equality laws.

The report made some recommendations, including **An Accountability for Algorithms Act** that establishes a corporate and public sector duty to undertake, disclose and act on pre-emptive Algorithmic Impact Assessments (AIA), which includes a dedicated equality impact assessment.

On 25 November 2021, all 193 Member States of the United Nations Educational, Scientific and Cultural Organisation (UNESCO) unanimously agreed to a Recommendation on the Ethics of Artificial Intelligence. To quote the UNESCO headline, this is "the first-ever global agreement on the Ethics of Artificial Intelligence". The text[10] itself is admirable, but the Recommendation is only voluntary, so it is not a binding or enforceable agreement.

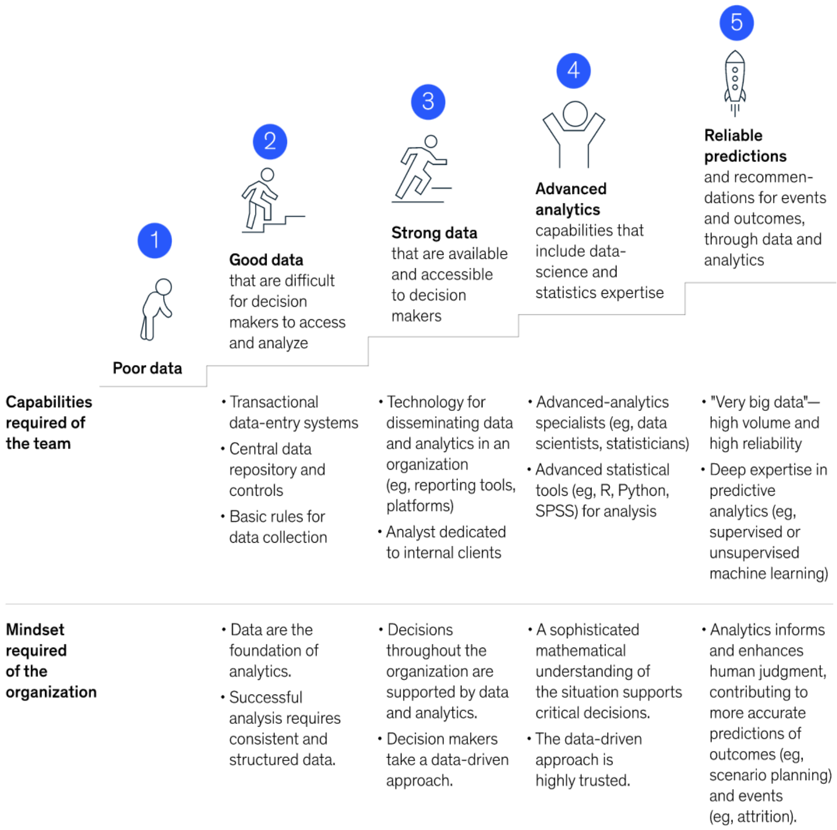
[5] https://www.linkedin.com/business/talent/blog/talent-strategy/future-of-recruiting
[6] https://business.linkedin.com/content/dam/me/business/en-us/talent-solutions/talent-intelligence/workforce/pdfs/Final_EMEA_Rise-of-Analytics-Report.pdf
[7] A recording of the parliamentary session on 23 June 2021 can be found on YouTube: UK National AI Strategy 2021 (Ethics & Governance) - YouTube
[8] APPG on the Future of Work (futureworkappg.org.uk) and Future of Work APPG (parallelparliament.co.uk)
[9] The New Frontier: Artificial Intelligence at Work (v7-08.11.2021) (squarespace.com)
[10] https://en.unesco.org/artificial-intelligence/ethics#recommendation

Despite that, it is obvious that we are firmly on our journey to embrace ethical AI - fair, inclusive, compliant and can be trusted. The paper sets out some best practice examples we have encountered for the readers to consider, and highlight areas that require reflections and more research.

> "Debug your data first, not your programme"
> **– Fernanda Viégas, Principal Scientist, Google Brain**

According to McKinsey[11], most large organisations have built up people-analytics capabilities, with 70% of executives considering it as a top priority. Underlying its people analytics framework is the importance of quality data (Figure 1).

Figure 1: McKinsey 5-step stairway model of people analytics



To remove bias in human capital management, avoid prejudices, groupthink and clubbiness, we must use thoughtful, curated data with domain specific applications for recruitment and talent management. Professor Andrew Ng, a co-founder of Coursea, former head of Google Brain, and former Chief Scientist at Baidu, observes that 80% of AI developers' time is spent on data preparation.

His findings are echoed by researchers who have warned that the 'big data' movement has created the **false expectation where 'more data makes algorithmic prediction better'**. Quite the contrary, data quality has a downstream impact, as it is deemed "*the most undervalued and de-glamorised aspect of AI*". Data cascades[12], referred to compounding events like landslides, causing negative, downstream effects due to data issues - are triggered by conventional AI practices that undervalue data quality. Cascades occur because of the default assumption that datasets were reliable and representative of the underlying population we try to measure and make predictions in performance. Application-domain expertise is therefore crucial in determining the performance and

[11] https://www.mckinsey.com/business-functions/organization/our-insights/how-to-be-great-at-people-analytics
[12] https://research.google/pubs/pub49953/ and https://ai.googleblog.com/2021/06/data-cascades-in-machine-learning.html?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+blogspot%2FgJZg+%28Google+AI+Blog%29

reliability of domain specific AI analytics, and yet such expertise is not often sought out until outputs are delivered and questions are raised by domain experts when they conduct a common sense check.

In the context of human capital management, data that allegedly reflects ideal employee attributes is used for training algorithms to assess candidates' abilities in developing and perfecting those attributes.

For an AI model to generalise well, it needs to be trained on representative data reflective of real-world settings. One could argue that any human resources or talent assessment domain expertise can be factored into an AI model. However, as Dr. Katie Creel, an Embedded EthiCS fellow at Stanford University has highlighted, whilst individual companies or organisations rank job applicants by their own metrics, commercial AI algorithms increasingly are built on foundation models (Figure 2), such as BERT and GPT-3. They form the backbone of different commercial AI chatbots and sentiment analysis developed by third parties, including start-ups and entrepreneurs.

These foundation models homogenise assessment frameworks and standardise assessment outcomes. As the same assessment hierarchy or ranking is likely to be used across sectors and geographies, the human capital management AI can set rules that monopolise access to opportunities. This is the unintended consequences of homogenisation highlighted by researchers from the Human-centred AI (HAI) centre at Stanford University[13].

Figure 2: Foundation Models

The type of AI used matters. There are two main models of AI that may be used for machine learning to make predictions using workplace data; static or dynamic[14]. Static models train using historical workplace data to develop the algorithm to be used as the basis for predicting how a new sample of employee data lines up to workplace outcomes. Once the model is trained on historical data, the algorithm remains static throughout its use in future predictions. Due to the fast paced nature of change in the workplace, especially seen during the pandemic with rapid shifts in how we work, this model lacks the flexibility to adapt to changing data. A more suitable type of AI to use in human capital management is a dynamic model. With this type of AI, new data is continually entering the system and tweaking how the algorithm works. Through the incorporation of new data into the algorithm, these types of predictions will be able to adapt to changes at work. If a company uses a vendor for customising and updating their models, it is important to consider if becoming reliant on suppliers could become a risk.

Throughout this paper, we take a human-centric approach to evaluate the different pillars of AI applications in human capital management. For investors that are familiar with the United Nations Guiding Principles for Business and

[13] A copy of the paper 'On the Opportunities and Risks of Foundation Models' can be found here: https://arxiv.org/pdf/2108.07258.pdf
[14] Zhang, D., & Wang, J. (2021). Recommendation Fairness: From Static to Dynamic. *arXiv preprint arXiv:2109.03150.*

Human Rights (UNGP), the idea of human rights goes beyond physical experience to include the right to social security, to education and water; human rights encompass humans having the right to be treated with dignity and respect[15]. Therefore, we consider the mental health impact of workplace surveillance, whilst acknowledging the need that some form of monitoring is necessary for certain roles for compliance and risk management. Other forms of monitoring are beneficial to both the employer (when measuring performance) and the employee (when targeting self-improvement).

In all AI people analytics, the foundation should be built on individual consent and transparency that the employees, as well as their managers and business division leaders, are fully aware of the objectives of using people analytics, the extent in which they are used, and that the outputs of analytics are subject to the scrutiny of employees being measured and managers who use them to support decision-making.

Humans are the users of AI analytics, not a product of their models and algorithms.

> "Make it so."
>
> **– Jean Luc Picard.**

Definitions

Defining Artificial Intelligence ("AI")

Artificial intelligence is the simulation of human intelligence processes by machines. Specific applications of AI include expert systems, natural language processing, speech recognition and computer vision. A more detailed explanation of the definition of the term can be found in Background section of the paper Investors Expectations on Responsible Artificial Intelligence and Data Governance (2019)[16].

The UNESCO Recommendation on the Ethics of Artificial Intelligence[17] defined AI systems as '*information-processing technologies that integrate models and algorithms that produce a capacity to learn and to perform cognitive tasks leading to outcomes such as prediction and decision-making in material and virtual environments.*'

Defining Human Capital Management

Human capital management (HCM) is a set of practices related to people resource management. According to Gartner, a global technology consulting firm, these practices are focused on the organisational need to provide specific competencies and are implemented in three categories: workforce acquisition, workforce management and workforce optimisation.

Core strategic activities in human capital management therefore includes:

1. Workforce acquisition
   1.1. Recruitment
   1.2. On-boarding
   1.3. Mobility management: cross departmental and/or overseas postings
2. Workforce management
   2.1. Competency management: cataloguing, managing, and developing skills
   2.2. Remuneration strategy
   2.3. Team dynamics assessment and monitoring

[15] https://www.ungpreporting.org/resources/the-ungps/
[16] https://www.hermes-investment.com/uki/wp-content/uploads/2019/04/investors%E2%80%99-expectations-on-responsible-artificial-intelligence-and-data-governance.pdf
[17] https://unesdoc.unesco.org/ark:/48223/pf0000379920.page=14. Annex page 4.

3. Workforce optimisation
    3.1. Workforce planning
    3.2. Performance and productivity management
    3.3. Learning (education and training)
    3.4. Disciplinary

LinkedIn Learning has an Applied AI for Human Resources course[18], which covers five specific use cases:

- Talent acquisition;
- Self-service help;
- Organisational analysis and design;
- Training and development; and
- Attrition management

We categorised the above topics into three chapters (Figure 3): Hiring, Culture, Performance. The overarching task in ethical AI for human capital management however, is to begin with Governance (Chapter 1).

Figure 3: Categorisation of the Human Capital Management Process

Hiring	**Culture**	**Performance**
Workforce planning	Competency	Remuneration
Recruitment	Team collaboration	Attrition
On-boarding	Mobility	Learning

[18] https://www.linkedin.com/learning/applied-ai-for-human-resources

Chapter 1. Governance

The decision to embed AI technologies in certain business and operational functions are often driven by frontline employees as a business need request. Whilst they are well placed to identify such a need, the rapid development and deployment of new technologies requires a comprehensive AI embedded strategy and risk management process, led by the C-level executives with board oversight, especially when such technologies are involved in tasks that have significant human and business impact, such as interactions related to customers, suppliers, third-party supply chains, and employees. (For examples from the UK regulated financial services industry and suggestions that, in our view, apply more generally, see paragraph 2.1.4 in the Legal and Regulatory Considerations chapter.)

While in our view good corporate governance requires board consideration and oversight of the decision to adopt AI in the workplace and, more widely, in customer and supplier interactions, there are sometimes even more compelling legal and regulatory considerations that will require or, at the very least, persuade corporate boards and senior management to maintain sufficient oversight of their companies' deployment of technologies and related third party services, including many forms of outsourcing, in each case that would have a significant internal or external impact on their businesses. (For examples from the UK regulated financial services sector, see paragraphs 2.1.3 – 2.1.5 in the Legal and Regulatory Considerations chapter.)

We observe in the Legal and Regulatory Considerations chapter that there is very little current law and regulation around the world that apply specifically and directly in the context of the introduction and use of AI in the workplace. This statement applies equally to the more general issue of corporate governance in this context. However, current law, regulation and (where applicable) regulatory expectations of corporate boards or governing bodies offer good starting points for governance and oversight considerations in the context of introducing AI into internal and external operations; and, in the case of current law and regulation, may be expected to apply through legal challenges and judicial decision-making in relation to the introduction and use of AI systems and processes in the workplace, in supplier/service provider arrangements (including outsourcing), and through third-party supply chains.

The Financial Stability Institute (FSI) of the Bank for International Settlements (BIS), in its FSI Insights on policy implementation No 35, *Humans keeping AI in check – emerging regulatory expectations in the financial sector*,[19] makes the point that, in that sector, existing requirements relating to governance, risk management, and the development and operation of existing models apply equally to AI models, and include "governance requirements placing responsibility on the use of such models on boards and senior management of financial institutions"[20]. The FSI goes on to say that financial services supervisors expect AI models to be transparent, not only to comply with risk management principles, but also to enable supervisory scrutiny of those models.

With regard to existing corporate governance standards that require general accountability requirements for banks and insurance companies, the FSI refers to Basel Core Principles (BCP) 14 for banks, Insurance Core Principles (ICP) 7 for insurers, and the Basel Committee on Banking Supervision (BCBS) principles on corporate governance (BCBS (2015)). The FSI refers specifically to BCP 15 and ICP 17, which apply in relation to the use of risk models, and "assigns [sic] ultimate responsibility to the board and senior management and for them to understand the consequences and limitations of model outputs"[21], which we assume would apply in relation to the introduction and ongoing deployment of AI models in those regulated sectors, both in internal (including the workplace) and external

[19] Authored by Jermy Prenio and Jeffery Yong, published in August 2021, at: https://www.bis.org/fsi/publ/insights35.pdf.
[20] FSI Insights on policy implementation No 35, as cited above, at page 1.
[21] See FSI Insights on policy implementation No 35, as cited above, paragraph 24 at page 12.

operations, and certainly where the deployment has or would have a material impact on those operations, the conduct of business, impact on customers, and on operational resilience more generally.

Human capital management is one area where AI analytics applications have experienced explosive growth in the past few years, which accelerated since the COVID-19 pandemic began, as more business activities take place digitally. Digital activities, including video recordings of job interviews, contribute to a library of workplace activity data. It is debatable whether regulations concerning fair employment practices are more or less stringent than those applying to financial services for retail customers. Nonetheless, companies that seek to improve diversity, equity and inclusion (DEI), workforce productivity and employee experiences have the responsibility to address the risks presented by AI technologies when exploring the opportunities they bring.

The **Objectives** of deploying AI technologies in human capital management should be about demonstrating a firm's commitment towards trusted AI principles with actions to establish responsible and fair practices for key stakeholders, including shareholders, employees, customers, regulators, suppliers and civil society.

Recommendation 1: Clarify Intentionality

Companies should define the **scope of AI applications** in the domain of human capital management. The scope should reflect:

Why technology enables fairer, more efficient, transparent and inclusive practices;

How the technology is expected to deliver its promises, and

What happens if it fails? Is this added to the risk register and disclosed in annual report as material risks where appropriate?

The scope should consider intentional and unintentional consequences, and methods to mitigate identified unfavourable outcomes. Let's consider using AI to measure productivity as an example.

Rationale 1: The risk of gaming the system - if Microsoft Productivity Scores[22] were included as a key performance indicator (KPI) for measuring employee performance, which has an impact on promotion opportunities, bonus allocation and remuneration increases, employees would be interested in understanding how to change their behaviour or at least the activities that indicate their change in behaviour, to maximise points to their advantage. This is no different from a company trying to understand how third party service providers put together sustainability scores or criteria for indices inclusion in order to maximise their scores or their chances to be included in a particular index to tap additional capital.

According to the OECD[23], productivity is *commonly defined as a ratio between the output volume and the volume of inputs*. In other words, it measures how efficiently production inputs, such as labour and capital, are being used in an economy to produce a given level of output.

Microsoft productivity scores are composed of eight factors, ranging from content collaboration, network connectivity to how often applications are updated, the number of online meetings and user applications experience. Factors are weighted equally across all factor domains. Companies can benchmark against peers, and internally, teams versus teams within an organisation.

This means that Microsoft productivity scores measure the **productivity of the use of their software, not the productivity of the workforce using it.** For example, a higher number of meetings does not represent better employee performance even if one may receive a higher score. It represents higher utilisation of Microsoft teams, which in turn justifies the acquisition cost of the product. Note that fatigue of over using teams chats, emails,

[22] https://docs.microsoft.com/en-us/microsoft-365/admin/productivity/productivity-score?view=o365-worldwide
[23] https://www.oecd.org/sdd/productivity-stats/40526851.pdf

meetings could be deemed the 'deadly sins of digital work habits'[24]. As such, the highest utilisation may not be indicative of higher employee performance.

Insights 1: Understand that words have different meaning in the context of use.

A Facebook or Instagram friend is a category of 'Friends', but it does not necessarily represent all friends. For some, social media friends are their followers and contacts; they are not 'friends'. For others, they could be a mix of friends, contacts and followers; the proportion of the mix varies by individuals.

All terms and attributes/factors used in AI Analytics need to be clearly defined, including their reference framework that provides the relevant business context. For a CEO, measuring their productivity by the number of online meetings they organise and attend seemed totally unintelligent and guarantee a disaster for the business. On the other hand, for a call centre worker, counting the number of calls where customer complaints are satisfactorily resolved given the amount of time and resources spent, is a valid measure of productivity.

Even so, taking a human-centric approach, the time frame of input into this productivity measure should monitor spikes in under- or over-performance that may have been impacted by personal circumstances, such as a close family member passed away from COVID-19.

[24] https://www.swoopanalytics.com/blog/digital-behavioural-change-are-you-motivated-enough

EXPERT OPINION – Intentionality Explained: Adding Value Beyond the Learned Model – by Dr Juergen Rahmel, Head of Digital Technology and Innovation, Compliance APAC, HSBC

Intentionality is created by asserting the proper functioning of the AI system and setting applicability boundaries and out-of-scope situations.

"AI is not plug and play."

When applying AI to a particular use case, it is important to separate the use of AI into two steps: step 1: learning; step 2: decision-making based on the learned model. Solution finding is any activity that an AI system typically is expected to achieve. These can be categorization/classification, interpolation/regression, criterion optimization, planning and related capabilities.

In step 1, the attempt is to provide the best possible foundation for learning from previous experiences, labelled with previous outcomes. Learning should take place with high quality and appropriate transparency, based on data that reflects the real and correct characteristics of the use case in question. The selection of learning algorithm and representation of knowledge should match the nature of data and the expected logic of decision making needed in step 2. The data is intended to be 'bias-free' **with respect to the planned overall capability of the applied AI solution**. Carefully choosing a representative data set is key in this step.

Step 2 can be seen as logically a separate activity, bridging the gap between the learned model and the expected decision making capability of the system. Step 2 will be able to contain forward looking criteria, which may not be embedded into the mathematical function of the trained model of the AI system actually represents, in order to be able to add value beyond the learned model.

The validity of the system composed of steps 1 and 2 should then be verified through a macro-feedback-loop on a governance level. This is aiming to answer the master-question: is this particular system of data, learning and decision making 'in principle' capable of providing a solution to the use case that the (future) users will have in mind? That is a question to be posed at conception stage and constantly upheld during each instance of applying the AI.

Illustrative example: imagine you are NBA Basketball coach of an ambitious team, trying to use a trained AI system to assist with the evaluation and selection of players. It is the year 1988 and any model trained on current data will tell you that players on the Point Guard position should on average be circa 6 foot 2 (or 187 cm) of height, with a relatively small variation. This is based on representative data of the current and previous seasons.

The decision making of a coach however should be able to include other factors, such as e.g. mental strength and perseverance needed for a tough sport on highest level of professional requirements and also the coach's favourite attack and defence models. It was this wider structure of decision-making criteria – beyond any statistical model relying on the past - that led this coach in 1988 to call Muggsy Bogues into the team – a player who, when compared on the height criterion with 5 foot 3 (or 160cm) and other purely physical measures, would have been perceived as outlier without a chance. Bogues' many years of successful participation in the NBA speak for the coach's decision to include more than only 'objective' criteria coded into a statistical model.

Recommendation 2: Establish Processes

Companies should establish **clear processes of AI applications on-boarding and monitoring** in the domain of human capital management. The processes should reflect:

1. **What** does the technology do that is expected to bring better outcomes than the current process? What is the verification and audit process?
2. What **inputs are required** to customise the application that will meet the specific needs of the company?
3. What **monitoring and changes** will ensure it continues to be relevant and fit for purpose? For example, assessing risk levels and conducting scenario analysis where appropriate.

The more complex the task(s), the more sophisticated the key performance indicators (KPIs) of success should too be constructed 'backward-looking' based on the desired outcomes. In documentation, independent sources of information should be sought to support a thorough due diligence process. As highlighted by the UN High Commissioner for Human Rights Michelle Bachelet in a recent report, **businesses have often rushed to incorporate AI applications, failing to carry out due diligence**[25].

Companies should form an assessment team and plan for the long term, such as how will test for bias be conducted before and after deployment; define assessment methods, understand workflows involved, set target outcomes and success criteria for each technology application. The workflow for the assessment of AI analytics developed by third party Fintech, Regtech, or HRtech providers may be more complicated than analytics developed in-house using proprietary data only.

Rationale 2: A robust process is necessary for foreseeable and unforeseen risks. Companies should have well documented board meeting notes to support audits and any potential investigation. Companies should undertake adequate risk assessment, review or adequately plan for ongoing monitoring before the commencement of any offshore or outsourcing arrangements to ensure that trusted party would deliver products and services in compliance with associated regulatory obligations.

Insights 2: Maintain an updated risk registrar for each technology application.

Taking a human-centric approach, mindful that the myth of **'what gets measured gets managed' could become the management of KPIs instead of the underlying issues.** A periodic review by a diverse panel of experts should be put in place to ensure that the status quo is fit for purpose.

Data cascades were set off by a lack of documentation across various cross-organisational relations (within the organisation, with field partner organisations and data collectors, and with external sources).

Using data where bias considerations has been thoroughly analysed and adjusted should always be the first choice when using AI in human capital management. However, when companies do not have access to completely unbiased data, there are methods for altering the algorithm to result in more fair outcomes. This can be done by using metaheuristic algorithms which use optimisation to solve problems[26]. In the case of bias, the optimisation would target diversity to amend the bias and validity to ensure the algorithm is still making the most accurate decisions possible.

In the evaluation process, there should also be a reference point with which to compare the outcomes of AI decisions versus that of human decision making. This way, we can see whether AI makes better or worse decisions.

[25] https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=27469&LangID=E

[26] Abdel-Basset, M., Abdel-Fatah, L., & Sangaiah, A. K. (2018). Metaheuristic algorithms: A comprehensive review. *Computational intelligence for multimedia big data on the cloud with engineering applications*, 185-231.

EXPERT OPINION – On Correcting for Bias – by Professor Elias Aboujaoude MD, Stanford University

Data centric AI advocates highlight that it is possible to re-engineer or modify data if bias is discovered[27]. Consistency of labelling data and a systematic way of going about cleaning it and correcting errors improve data quality. This approach is applied in other disciplines, such as medical research.

As noted by Professor Elias Aboujaoude MD, "It is not unusual in medicine to build entire treatment algorithms based on "open label" data or data from individual cases that later proves to be flawed. This foundational data can, and should, be re-examined. In clinical medicine, we do that by using the gold standard of the double-blind, placebo-controlled study. If these much better, less biased controlled studies yield different results, the entire algorithm is revisited. Artificial Intelligence systems would benefit from such a re-examination of their foundational data to check for any flaws and biases—and to change course and be willing to go back to the drawing board when flaws and biases are identified."

Recommendation 3: Ensure Accountability

Companies should have a governance structure with board level oversight for ethical AI applications in business and operations. This is due to the potential human impact, including behavioural influence that AI-empowered people analytics may have on individuals. The oversight should reflect:

1. **Who** would have the skills, expertise and experience to provide input, judgement and oversight?
2. **What** content should be included in the relevant materials for review and approval?
3. **What** additional due diligence may be necessary and what resources would the company provide to board members on material issues arising from the use of AI people analytics?

Companies must understand how vendor interprets fairness and bias, and how much customisation is needed to suit specific needs, and the internal and external resources that may be required to achieve the expectations.

Rationale 3: AI presents new complexities, as the difference between right and wrong may be more nuanced, and subject to interpretation and judgement. Ethical committees may present potential solutions and options, but ultimately it will be the board's decision to decide and mandate the ethical course.

Companies should make a commitment to ethical impact assessment audits of their AI systems to predict consequences, mitigate risks, avoid harmful consequences, and facilitate participation. This could include involving different stakeholders in their AI governance and adding the role of an independent AI Ethics Officer at the responsible committee(s).

Insights 3: Ensure that the internal algorithmic audit team has the right mix of business experience and technical expertise. Where appropriate, seek external assurance and third party audits.

EXPERT OPINION – On Privacy – by Anna McDonald, Secretary, The Church of England Ethical Investment Advisory Group

"Privacy is essential to the protection of our human dignity and human autonomy. It is personal – but it is also collective – when one person's privacy is exposed – the dignity and agency of others is impacted. Data for AI systems should be collected, used, shared, archived and deleted in ways that respect this right to privacy and are consistent with international law and internationally agreed global norms such as the UNGPs."

[27] Data-centric AI: Real World Approaches, 11 August 2021. https://www.youtube.com/watch?v=Yqj7Kyjznh4

Recommendation 4: Mandate Transparency

Transparency is closely linked to companies' legal and ethical obligations to respect human rights, for example as set out in the UN Guiding Principles on Business and Human Rights (UNGPs) and the UN Global Compact. It is also cited in the UNESCO Recommendation on the Ethics of Artificial Intelligence, which noted that transparency is "often a crucial precondition to ensure that fundamental human rights and ethical principles are respected, protected and promoted necessary for relevant national and international liability legislation to work effectively".

Companies should mandate sufficient disclosure and adequate transparency in the use of AI in human capital management. The transparency standard should reflect:

1. Objective of use, data used and reference framework for contextual application
2. Accountability structure, with a thorough understanding of how users interact with the systems that might lead to the risks of disparate impact;
3. Monitoring and reporting requirements

What is Disparate Impact?

Disparate impact in United States labour law refers to practices in employment, housing, and other areas that adversely affect one group of people of a protected characteristic more than another, even though rules applied by employers or landlords are formally neutral. Disparate impact focuses on 'outcomes fairness', but there is also a need to ensure 'implementation fairness' in how AI systems are implemented and deployed and the necessity for companies to adopt a process of continual reflection and verification of vendors.

Rationale 4: The objective of using AI analytics for certain tasks need to be well-argued, the limitations assessed and the due diligence process properly documented and followed to ensure sufficient disclosure and adequate transparency.

Insights 4: Transparency impacts the confidence of job candidates and employees regarding AI, especially when they have become the subject of assessment. Pushback could occur when individuals are resistant to change, or have doubts about the fairness of the assessment. As AI is still relatively new and is a complex subject, when engaging with employees on proposed new practices, it is important to make explanations clear and concise, to prioritise opt-in rather than opt-out participation if it is not business critical or compliance mandated, and to take into considerations the concerns of different employee resource groups. There is also nuance in how transparency will impact the perceptions of AI.

For example, when looking at fairness, including additional information on how AI operates, could hinder fairness ratings[28]. When information is given on how AI works, but it is not sufficiently explained, could deter employees through creating confusion. Hence, it is important for the AI language to be translated in a way that those who are not experts nor familiar with the topic can still understand the key messages. There needs to be open channels available for clarification, either in groups or individually.

Would an Opt-in led Approach lead to Under-representation?

An opt-in led approach ensures consent, but it is important to note that it is not a silver bullet. For example, an individual with a disability might choose to opt out of an AI assessment or not to contribute data to train an algorithm because they may be concerned that disclosure of their disability to an employer might constrain their access to healthcare provided or funded by that employer. In this case, the training data may exclude the appropriate representation of a certain subgroup within the workforce. A company is therefore recommended to have Employee Resource Group (ERG) to debate the potential impact of such exclusion and work with the data scientists team and relevant product vendors to minimise risk of non-representation.

[28] Newman, D. T., Fast, N. J., & Harmon, D. J. (2020). When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. Organizational Behavior and Human Decision Processes, 160, 149-167.

The UK Cabinet Office's Central Digital and Data Office ("CDDO") has published an algorithmic transparency standard, composed of data standard, application template and guidance. The data standard is composed of over 30 use-case attributes capturing information from responsible owner, supplier, access terms, benefit, model type, frequency of use, maintenance procedure, system architecture, human decisions involvement to training, impact assessment, risk identification and mitigation details (Figure 4). Although the tools are targeted at government institutions and public sector bodies, they are useful materials for private sector institutions committed to algorithmic transparency.

Figure 4: UK Algorithmic Transparency Data Standard

Guidance

Algorithmic transparency data standard

Updated 2 December 2021
Download CSV 9.22 KB

Reference number from the word template	attribute (an asterisk indicates a multi-value field)	name	tier	category	type	description	required
Not set	name	Name	Tier 1	Overview	UTF-8 string	The attribute 'name' is the colloquial name used to identify the algorithmic tool.	TRUE
Not set	description	Description	Tier 1	Overview	UTF-8 string	The attribute 'description' is a description to give a basic overview of the purpose of the algorithmic tool. It should include: - how the algorithmic tool is used - why the algorithmic tool is being used	TRUE
Not set	website	URL of the website	Tier 1	Overview	URL	The attribute 'website' is the URL reference to a page with further information about the algorithmic tool and its use. This facilitates users searching more in-depth information about the practical use or technical details.This could, for instance, be a local government page, a link to a GitHub repository or a departmental landing page with additional information.	FALSE
Not set	contact_email	Contact email	Tier 1	Owner and responsibility	UTF-8 string	The attribute 'contact_email' is the email address of the organisation, team or contact person for this entry.	TRUE
1.1	organisation	Organisation/ department	Tier 2	Owner and responsibility	UTF-8 string	The attribute 'organisation' is the full name of the organisation, department or public sector body that carries responsibility for use of the algorithmic tool. For example, 'Department for Transport'.	TRUE
1.2	team	Team	Tier 2	Owner and responsibility	UTF-8 string	The attribute 'team' is the full name of the team that carries responsibility for use of the algorithmic tool.	Not set
1.3	senior_responsible_owner	Senior responsible owner	Tier 2	Owner and responsibility	UTF-8 string	The attribute 'senior_responsible_owner' is the role title of the senior responsible owner for the algorithmic tool.	Not set
1.4	external_supplier*	Supplier or developer of the algorithmic tool	Tier 2	Owner and responsibility	UTF-8 string	The attribute 'supplier' gives the name of any external organisation or person that has been contracted to develop the whole or parts of or the algorithmic tool.	TRUE
1.5	external_supplier_identifier*	External supplier identifier	Tier 2	Owner and responsibility	UTF-8 string	The attribute 'external_supplier_identifier' gives, if available, the Companies House number of the external organisation that has been contracted to develop the whole or parts of or the algorithmic tool. You can get a company's Companies House number by finding company informationor using the	FALSE

BrainCo, a startup from Harvard Innovation Lab pivoted its cognitive training technology from students productivity surveillance to stress relief, mental wellness and cognitive training[29]. It shows that the same technology can be used for different purposes. If the objectives, methodologies and functionality of the technology is properly disclosed and users can choose to sign up to using it, workplace surveillance could transform into productivity assistance programmes.

Companies can use the World Economic Forum (WEF) Human-centred AI for Human Resources toolkit[30] to establish the basic governance framework that meets the four recommendations highlighted in this chapter:

- Clarify intentionality – determine the purpose of adopting the AI-based tool, and how the technology is expected to deliver its promises;
- Establish processes – form an assessment team and plan for the long term, such as how the test for bias will be conducted before and after deployment; periodically assess model drift;
- Ensure accountability – consider who should be brought in on the decision to adopt the AI tool and to monitor it; understand how the vendor interprets fairness and bias, as well as how much customisation is needed to suit specific needs if vendors are used; and
- Mandate transparency – ensure explainability and understand how users interact with the AI tool.

> "AI requires us to be explicit about the future and it is the interrogation process that we as human beings need to confront."
>
> **– Stuart Russell, Professor of Computer Science, University of California Berkeley**

[29] https://www.independent.co.uk/news/world/asia/china-schools-scan-brains-concentration-headbands-children-brainco-focus-a8728951.html, https://www.yicaiglobal.com/news/brainco-boss-says-neurotherapy-headband-is-for-education-not-surveillance and https://brainco.tech/
[30] https://www.weforum.org/reports/human-centred-ai-for-hr-state-of-play-and-the-path-ahead/

Chapter 2. Legal and Regulatory Considerations

This legal and regulatory considerations chapter and other legal and regulatory inputs in this paper are intended to provide general information about certain current, recent and anticipated developments which may be of interest to readers. This chapter and those inputs are not intended to be comprehensive nor to provide any specific legal advice and should not be acted or relied upon as doing so. While this chapter and those inputs also refer to legal and regulatory developments in jurisdictions other than England and Wales, Macfarlanes LLP is not qualified to advise in relation to those jurisdictions and those references should be read as no more than reported statements of such developments and the law or regulation in the jurisdictions concerned, and not as authoritative statements of any law or regulation. Professional advice appropriate to the specific situation should always be obtained. Macfarlanes LLP is a limited liability partnership registered in England with number OC334406. Its registered office and principal place of business are at 20 Cursitor Street, London EC4A 1LT. The firm is not authorised under the Financial Services and Markets Act 2000 - but is able in certain circumstances to offer a limited range of investment services to clients because it is authorised and regulated by the Solicitors Regulation Authority. It can provide these investment services if they are an incidental part of the professional services it has been engaged to provide.

"Artificial intelligence and algorithmic decision-making tools have great potential to improve our lives, including in the area of employment. At the same time, the EEOC is keenly aware that these tools may mask and perpetuate bias or create new discriminatory barriers to jobs. We must work to ensure that these new technologies do not become a high-tech pathway to discrimination."

- Charlotte A. Burrows, Chair of the U.S. Equal Employment Opportunity Commission (EEOC)[31], 28 October 2021

2.1. Introduction

- 2.1.1 It has been widely accepted that corporate legal and regulatory compliance goes beyond ethics as norms, adherence to ethics, and investors' expectations of such adherence. The consequences of corporate failure to comply with law and regulation may include individual and corporate criminal and civil liability to varying degrees and in varying amounts[32], as well as irremediable reputational damage among customers, employees, suppliers, shareholders, and other direct or indirect investors. And failure to comply with law and regulation may also threaten the very existence of a company, in some cases leading to insolvency, regulatory intervention and closure, or other forms of winding-up. It is a truism that, at the very least, investors expect companies in which they hold direct or indirect investments to comply with law and regulation wherever those companies operate.

[31] Published in a press release of the EEOC on 28 October 2021, entitled The U.S. Equal Employment Opportunity Commission (EEOC) is launching an initiative to ensure that artificial intelligence (AI) and other emerging tools used in hiring and other employment decisions comply with federal civil rights laws that the agency enforces: https://www.eeoc.gov/newsroom/eeoc-launches-initiative-artificial-intelligence-and-algorithmic-fairness
[32] For example, the proposed EU Artificial Intelligence Act (AIA) would impose maximum fines for the most serious infringements of the AIA of either 30,000,000 EUR or 6% of total worldwide annual turnover for the preceding financial year, whichever is higher: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206

- 2.1.2 As such, legal and regulatory compliance, and the risks and exposure resulting from non-compliance, are and must be vital considerations for the board and senior management of any company operating under generally accepted corporate rules, norms, and values.

- 2.1.3 In the UK, the Senior Managers and Certification Regime (SM&CR) applies, though to differing degrees, to all financial services firms regulated by the UK Financial Conduct Authority (FCA). Senior managers in regulated firms are under a statutory "duty of responsibility". If a regulated firm infringes a FCA rule, the senior manager responsible for that area concerned could be held personally accountable if they did not take reasonable steps to prevent or stop the infringement[33]. Currently, only the largest and most systemically important firms have a designated chief operations function under SM&CR (Senior Management Function 24), which includes responsibility for the internal operations and technology of the firm.[34] It is fair to say that, currently, in the technology field, the FCA's main focus is on technology systems (which include algorithmic trading and AI and machine learning (ML)) and operations that may impact the operational resilience and customer-facing conduct of firms, rather than, say, their internal HR systems and processes. But the signs are that this may well change.

- 2.1.4 Going back at least to its 2017/18 Business Plan[35], the FCA has been concerned that senior management in regulated firms understands how AI and ML are being deployed in them. On 1 August 2019, the FCA published in its *Insight Series* an article by Magnus Falk, entitled *The advent of AI is not just a matter for the technicians, those at the very top of firms must take responsibility for the big issues.*[36] The author makes the following key points which, although they are made in a UK context, in our view have wider significance for all boards and governance bodies, and not just in the financial services sector. The following points made by Falk are therefore worthy of consideration by all boards, everywhere.

  - The boards of regulated financial services firms currently oversee technology and operations, but AI systems and processes[37] present a step change, which will demand new skills and new areas of focus at board level. AI systems and processes are relatively new, and certainly new to many companies that deploy them.

  - In ethics and ethical conduct, AI presents new complexities, as the difference between right and wrong may be more nuanced, requiring a balanced view between business outcomes and ethical conduct. Ethics committees may present one solution, but ultimately it will be the board's responsibility to decide and mandate the right ethical course.

  - Regarding the explainability of AI, the "black box" excuse will not sit well or at all with regulators or customers. The challenge is deciding the level(s) of "sufficient" explainability. Boards will ultimately have to decide this question. If they are unable to understand what an algorithm is doing in/for their company, they need the confidence and integrity to admit that, and to seek an explanation from those who can provide such an explanation, whether from in or outside the company.

  - On the issue of transparency in the deployment of AI systems and processes, Falk rightly notes that this is different from explainability. Customers should be told how and when AI is being used to conduct their business and make decisions and effect transactions concerning their interests. We would say - and do so elsewhere in this chapter and this paper - that the same principle of transparency should apply equally to staff and the deployment of AI systems and processes in the workplace. In our view, boards should consider determining the principles of the corporate approach to transparency, and mandate responsibility to the responsible executive(s) accordingly.

  - Risk, and exposure to claims and liability – boards must be informed and maintain awareness of how AI could affect their company's risk and exposure to additional claims arising from the use of AI systems and processes, whether internally or through third party or outsourced services. AI poses different and wider risks for those who deploy it and rely on its outputs. And we say that the same consideration should be given to AI systems and processes used in the workplace.

[33] https://www.fca.org.uk/firms/senior-managers-certification-regime
[34] https://www.handbook.fca.org.uk/handbook/SUP/10C/6B.html
[35] See *UK Regulatory cross-sector and sector priorities and risk outlook, Risk outlook, Technology and innovation, Key risks.*https://www.fca.org.uk/publication/business-plans/business-plan-2017-18.pdf.
[36] https://www.fca.org.uk/insight/print/artificial-intelligence-boardroom
[37] The references in this section to "AI systems and processes" are intended to cover some or all, as may be applicable, of the AI technologies and software, algorithms, data sets and other data elements (individually or together, "AI systems") as well as the processes associated with the implementation and deployment of AI systems. References in this section to "AI" include "AI systems and processes", where the context requires.

"… [D]iversity and inclusion are key priorities for the FCA and ones where we'll be taking action to raise standards… Diversity and inclusion are critically important to the FCA. Firms that are diverse and inclusive deliver better outcomes for shareholders, consumers and markets."

- Sheldon Mills, Executive Director, Consumers and Competition, Financial Conduct Authority, A regulatory perspective: measuring and assessing culture, now and in the future, the role of purpose and the importance of D&I, speech delivered at the Investment Association on 24 September 2021[38]

- 2.1.5 Now the FCA has started to prioritise diversity and inclusion in regulated firms. We can assume that, at least in the UK, regulatory scrutiny would extend to the deployment in those firms of AI systems and processes in a way that would run counter to hiring, developing, and retaining a more diverse and inclusive workforce.
- 2.1.6 The deployment of AI systems and processes in a corporate environment may seem a long way off from posing material, let alone existential, legal and regulatory risks and exposure of companies to substantial, in some cases unlimited, liability, especially where AI may be used in mostly internal operations and processes affecting current or prospective employees. From a legal and regulatory perspective, how are such risks and exposure likely to arise?
- 2.1.7 As we show elsewhere in this paper, increasingly, employers are deploying AI systems and associated processes at some, many, or all stages of human capital management - from pre-hiring and recruitment, through performance management, monitoring and surveillance in the workplace or while working from home, to decisions to take disciplinary action against employees and those terminating employment. If the use of AI in the workplace is growing, so are the risks and potential exposure to legal and regulatory claims that could result from use of and reliance on AI systems and processes, together with the data sets that underlie the outcomes of automated decision-making (ADM), when they deliver unethical, unfair and unlawful outcomes.
- 2.1.8 And what of the law and regulation in this area? Is there any? Put shortly, there is a very small, but growing, body of laws and regulations that apply directly to aspects of the deployment of, and reliance on, AI systems and associated processes in human capital management. And there is a much larger body of existing laws and regulations in many countries that, we are told and may assume, can and will apply indirectly - by their extension through judicial interpretation - to the use of AI systems and processes that have delivered unlawful outcomes. Increasingly, institutions like the European Commission and governments at various levels are looking to regulate AI, including in its deployment in the workplace. The proposed EU Artificial Intelligence Act (AIA)[39] is one such measure. It will have comprehensive and very substantial legal and regulatory effects well beyond the EU's borders, because of its proposed extra-territorial application and the "Brussels Effect".[40] In paragraphs 2.2 – 2.4 of this chapter, we offer a little more detail on the current and prospective applicable legal and regulatory landscapes, though given space constraints, our account is neither intended nor able to be comprehensive.
- 2.1.9 What are the likely triggers for, and origins of, legal challenges to employers using AI in human capital management? Among other factors, there are growing activist movements developing in several countries that seek to identify and address the wrongs that are and may be caused by AI systems and processes in and around the workplace.[41] This movement includes the UK's Trades Union Congress (TUC), which commissioned from leading UK employment and discrimination lawyers a comprehensive report (described by the authors as a legal opinion) on the legal implications of AI in the workplace, *Technology Managing People – the legal implications*[42], published in 2021. This report builds on the TUC's 2020 survey and report on the way that AI is being deployed

[38] https://www.fca.org.uk/news/speeches/regulatory-perspective-measuring-assessing-culture-diversity-inclusion.
[39] https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206
[40] "[T]he global power that the European Union is exercising through its legal institutions and standards (and sanction mechanism), successfully export[ing] that influence to the rest of the world.", Anu Bradford, *The Brussels Effect*, 107 NW U.L. Rev. (2012): https://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=1081&context=nulr
[41] For a more general, US-centric, account, see *The Movement to Hold AI Accountable Gains More Steam*, Wired, December 2, 2021: https://www.wired.com/story/movement-hold-ai-accountable-gains-steam/?utm_source=onsite-share&utm_medium=email&utm_campaign=onsite-share&utm_brand=wired
[42] Robin Allen QC and Dee Masters, the AI Law Consultancy, Cloisters Chambers, London: https://ai-lawhub.com/technology-managing-people-the-legal-implications/

and perceived in UK workplaces.[43] Whether or not one agrees with some or all of the purposes of, and conclusions drawn in, the TUC reports, they show the likely origins and direction of travel of legal challenges in the UK and possibly elsewhere[44] to the use of AI in human capital management, as well as possible future changes to EU and UK law in this area. And, though these reports emerge from the UK and its industrial relations landscape and legal systems, we suggest that both reports have wider importance for the discussion of the considerations this paper seeks to promote.

- 2.1.10 Also, government agencies, regulators and policy makers are concerned about the increasing use of AI in the workplace and the actual and potentially negative effects of such use. They are now subjecting that use and reliance on AI systems and processes in the workplace to greater scrutiny. For example, in October 2021, the US EEOC announced its *Initiative on Artificial Intelligence and Algorithmic Fairness*, which is intended to ensure that the use of AI by employers at all employment stages complies with US federal anti-discrimination law[45]. EEOC Commissioner Keith Sonderling has written and been quoted extensively on this subject. He has indicated that the EEOC may instigate "Commissioner charges", that is, EEOC-initiated investigations unconnected to discrimination charges, to ensure that US employers are not using AI unlawfully. Commissioner Sonderling is reported as saying that such charges would facilitate enforcement, because job applicants and employees are often unaware that they have been excluded from a role because of flawed or improperly designed AI software or bias.[46] In the UK, the Information Commissioner's Office (ICO), which is responsible for enforcing the UK's data protection and privacy laws, will be updating its existing employment practices code to take account of, among other issues, the increasing use of AI in the workplace.[47] And the UK Equality and Human Rights Commission (EHRC) is working on guidance on how UK equality legislation[48] would apply to the use of new technologies in the workplace, including AI systems and processes. In its draft Strategic Plan for 2022-25, the EHRC states that it will work with employers "to make sure that using artificial intelligence (AI) in recruitment does not embed biased decision-making in practice."[49].

- 2.1.11 The examples in paragraphs 2.1.9 and 2.1.10 above signal that:
  - the use of AI systems and processes in the workplace is of growing concern to, and is either being, or is now more likely to be, scrutinised by, government agencies, regulators, and policy makers; and
  - employers should note the possibility of new law and the greater likelihood of enforcement action and legal claims by those agencies, regulators, employees, unions, works councils, and other workforce representatives.

- 2.1.12 This, then, is an illustration of how and why legal and regulatory risks and exposure to legal claims and liability will arise in relation to the growing deployment of AI systems and processes in human capital management.

[43] *Technology managing people: The Worker Experience*, Trades Union Conference, November 2020: https://www.tuc.org.uk/sites/default/files/2020-11/Technology_Managing_People_Report_2020_AW_Optimised.pdf

[44] An example being the recent litigation that a group of Uber drivers brought against Uber under article 22 (automated individual decision-making and profiling) of the EU's General Data Protection Regulation (GDPR), which concerns ADM and the factual and legal basis for such decision-making, in this case by Uber's 'Mastermind' algorithm in relation to Uber drivers. The case was brought in Amsterdam by the law firm, Ekker Advocatuur, which has published an unofficial translation of the legal proceedings at: https://ekker.legal/en/2020/10/26/uber-drivers-challenge-dismissal-by-algorithm/

[45] See footnote 30.

[46] This report is contained in, and is attributable to, *Brave New World: The EEOC's Artificial Intelligence Initiative*, Jill Pedigo Hall, von Briesen & Roper S.C., December 14, 2021,The National Law Review: https://www.natlawreview.com/article/brave-new-world-eeoc-s-artificial-intelligence-initiative

[47] "We are planning to replace our existing guidance with a new, more user-friendly online resource with topic-specific areas. We want to make sure that our new guidance addresses the changes in data protection law, reflects the changes in the way employers use technology and interact with staff and meets the needs of the people who use our guidance products.": https://ico.org.uk/about-the-ico/ico-and-stakeholder-consultations/ico-call-for-views-on-employment-practices/ . The ICO conducted a public consultation on the new code that closed on 28 October 2021. As at 9 February 2022, the new code has not been issued, though the ICO has published a summary of all responses received, available through the link above.

[48] Principally, the UK's Equality Act 2010.

[49] The EHRC's draft Strategic Plan 2022-25 is downloadable from: https://www.equalityhumanrights.com/en/our-work/our-strategic-plan-2022-2025-have-your-say. This quotation is from page 14 of the draft Plan.

◆ 2.1.13 While international legal instruments and agreements[50] and statements and codes of principles and ethics[51] governing the development and use of AI are important in several aspects, not least as guides to the ethical and desirable use of AI or to the adoption of laws and regulations that mandate fair and ethical AI, in this section we focus on current and prospective nationally applicable and actually or potentially enforceable laws and regulations, rather than such agreements, statements and codes.

◆ 2.1.14 The next part of this chapter outlines a small, representative sample, from a variety of jurisdictions and legal systems, of:

- current laws and regulations that apply specifically and directly to the deployment of AI systems and processes in the workplace;
- prospective law and regulation that will or are likely to apply specifically and directly to such deployment; and
- current laws and regulations that we expect would apply to such deployment, but through judicial interpretation and application, that is, as a consequence of being tested in courts and tribunals.

◆ 2.1.15 The purpose of these examples is to alert readers, wherever they are based, to the possibility, in relation to the use of AI in the workplace:

- of the existence of new (or relatively new) laws and regulations specifically governing such use, which, being new or relatively new, may not be well-known, known at all, or understood;
- of prospective laws and regulations that will, or are likely to, apply specifically to such use, when they are enacted; and
- that current laws and regulations may apply in relation to such use in ways that may be unexpected, non-obvious, or even counter-intuitive to the non-expert (and possibly even to an expert), because the application of such laws and regulations has yet to be tested in courts or tribunals, but could well be held to apply through cases brought by employees, unions, works councils or other employee or worker representatives.

And in each case, this is so that readers can, before they deploy AI systems and processes in the workplace, understand, plan for, and manage the legal and regulatory implications, risks and exposure that may arise.

2.2. Examples of Current Laws and Regulations that apply specifically and directly to the Deployment of, and Reliance on, AI in the Workplace

◆ **2.2.1 New York State:** *An act to amend the civil rights law, in relation to electronic monitoring*, 2021[52:] requires all employers in New York State to give prior written notice to newly hired employees if employers intend to monitor or otherwise intercept employee emails, text messages, telephone conversations, Internet access or usage of an electronic device or system. Such notices only need to be given to employees hired on or after the Act comes into effect, but not to existing employees. **Comes into effect 7 May 2022.**

◆ **2.2.2 New York City:** *A Local Law to amend the administrative code of the city of New York, in relation to automated employment decision tools*, 2021[53]: prohibits the use of AI ADM[54] tools in hiring and promotion processes unless they have first been audited for bias by an independent auditor. The purpose of the audit is to ensure that the AI ADM tool does not lead to disparate outcomes based on race or gender. Employers deploying such tools in hiring and promotion processes will be required to notify candidates that AI is being used, and which

[50] For example, the OECD's Principles on Artificial Intelligence, formalised in the Recommendations of the Council on Artificial Intelligence: https://www.oecd.org/digital/artificial-intelligence/ ; and
UNESCO's *Recommendation on the Ethics of Artificial Intelligence*, announced on 24 November 2021: https://en.unesco.org/artificial-intelligence/ethics#recommendation, referred to elsewhere in this paper.

[51] For example, The Future of Life Institute's *Asilomar AI Principles*, published in 2017, and of which there are 23 principles designed to guide AI developers in creating responsible and beneficial AI: https://futureoflife.org/2017/08/11/ai-principles/. And there are many voluntarily adopted corporate codes of AI ethics, including those referred to elsewhere in this paper.

[52] https://www.nysenate.gov/legislation/bills/2021/s2628

[53] https://aboutblaw.com/0vz

[54] "§ 20-870 Definitions… "Automated employment decision tool" means any computational process, derived from machine learning, statistical modeling, data analytics, or artificial intelligence, that issues simplified output, including a score, classification, or recommendation, that is used to substantially assist or replace discretionary decision making for making employment decisions that impact natural persons…." "

job qualifications and characteristics the AI will review. Job applicants will have the right to opt out and request that a human review their application instead. **Comes into effect 1 January 2023.**

- **2.2.3 State of Illinois:** *Artificial Intelligence Video Interview Act* (AIVI Act), 2019[55], as amended in 2021[56]: imposes restrictions on employers in the State of Illinois using videotaped interviews for hiring, and requires employers: (i) to notify an applicant in advance that their interview(s) may be analysed by AI to consider the applicant's fitness for the position; (ii) to inform applicants about how the AI works and what general characteristics it will use to evaluate them; and (iii) to proceed only with the consent of the applicant to being evaluated by AI. There are further restrictions on the sharing of applicants' videos and requirements for the destruction of videos afterwards. The 2021 amendment requires employers who rely solely on AI analysis of video interviews to determine whether an applicant will be granted an in-person interview, to collect and report data concerning the race and ethnicity of those not granted an in-person interview, and such data of those who are hired. The data must be reported annually to the Illinois Department of Commerce and Economic Opportunity, which must analyse the data and report to the Governor and General Assembly annually whether the data discloses a racial bias in the use of AI. **The original AIVI Act came into effect on 1 January 2020.**

2.3. Example of Prospective Law and Regulation that will apply specifically and directly to the Deployment of, and Reliance on, AI in the Workplace

> "On Artificial Intelligence, trust is a must, not a nice to have. With these landmark rules, the EU is spearheading the development of new global norms to make sure AI can be trusted. By setting the standards, we can pave the way to ethical technology worldwide and ensure that the EU remains competitive along the way. Future-proof and innovation-friendly, our rules will intervene where strictly needed: when the safety and fundamental rights of EU citizens are at stake."
>
> **- Margrethe Vestager, Executive Vice-President for a Europe fit for the Digital Age, on the publication of the AIA on 21 April 2021**[57]

- 2.3.1 The pre-eminent example of such prospective law and regulation is the EU's proposed Artificial Intelligence Act (AIA).[58] As mentioned in paragraph 2.1.8 above, the AIA will have comprehensive and very substantial legal and regulatory effects generally, and specifically in the context of AI in the workplace. A full outline of the AIA is beyond the scope of this paper, but there are many accessible summaries of its provisions.[59]

- 2.3.2 In essence, the AIA takes a risk-based approach to AI systems and will regulate them according to the risks the AI poses under the following risk-based classification.

- **2.3.3 Unacceptable risk:** AI systems that present a clear threat to safety, livelihoods and human rights will, subject to limited exceptions, be prohibited, e.g., AI systems and processes that:
    - "have a significant potential to manipulate persons through subliminal techniques beyond their consciousness";
    - may "exploit the vulnerabilities of specific vulnerable groups such as… persons with disabilities in order to materially distort their behaviour in a manner that is likely to cause them or another person psychological or physical harm"; and

[55]https://www.ilga.gov/legislation/fulltext.asp?DocName=&SessionId=108&GA=101&DocTypeId=HB&DocNum=2557&GAID=15&LegID=&SpecSess=&Session=
[56]https://custom.statenet.com/public/resources.cgi?id=ID:bill:IL2021000H53&ciq=ncsl&client_md=cf812e17e7ae023eba694938c9628eea&mode=current_text

[57]https://ec.europa.eu/commission/presscorner/api/files/document/print/en/ip_21_1682/IP_21_1682_EN.pdf
[58] https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206
[59] Apart from the European Commission's own summary in the document referred to in footnote 57 see e.g., see *EU Artificial Intelligence Act: The European Approach to AI*, by Mauritz Kop, of Stanford Law School, and a lawyer based in The Netherlands: https://www-cdn.law.stanford.edu/wp-content/uploads/2021/09/2021-09-28-EU-Artificial-Intelligence-Act-The-European-Approach-to-AI.pdf#:~:text=The%20EU%20AI%20Act%20sets%20out%20horizontal%20rules,artificial%20intelligence%20rules%20that%20apply%20to%20all%20industries. For a critical (but constructively so) approach, see the UK Centre for Data Ethics and Innovation (CDEI) Blog, *The European Commission's Artificial Intelligence Act highlights the need for an effective AI assurance ecosystem*, posted by Cailean Osborne on 11 May 2021 https://cdei.blog.gov.uk/2021/05/11/the-european-commissions-artificial-intelligence-act-highlights-the-need-for-an-effective-ai-assurance-ecosystem/

- may facilitate "manipulative or exploitative practices affecting adults" where such AI may be "covered by the existing [EU] data protection, consumer protection and digital service legislation that guarantee that natural persons are properly informed and have free choice not to be subject to profiling or other practices that might affect their behaviour".[60]

◆ 2.3.4 Although the AIA is subject to the EU's legislative process, and may well be amended over the coming months, we mention these potentially prohibited AI systems and processes to illustrate how, given the way AI systems and processes as mentioned elsewhere in this paper are being developed and deployed in the workplace, it is quite possible, if not probable, that some AI used in hiring, (e.g. interviewing) and in the operational workplace environment (e.g. monitoring and surveillance) may be classified as presenting an unacceptable risk and banned from use in the EU, in other words, as presenting an even higher risk than the specific, workplace-related, high-risk AI systems mentioned in paragraph 2.3.5 below.

◆ **2.3.5 High-risk:** AI systems that pose a high risk to the health and safety or fundamental rights of natural persons. Such systems will be allowed in the EU market subject to compliance with specified mandatory requirements, including prior conformity assessments (see below). Examples of such high-risk AI include a category directly relevant to the subject of this paper:

- AI systems intended to be used for recruitment or selection of natural persons, notably for advertising vacancies, screening or filtering applications, evaluating candidates in the course of interviews or tests;
- AI intended to be used for making decisions on promotion and termination of work-related contractual relationships, for task allocation and for monitoring and evaluating performance and behaviour of persons in such relationships".[61]

◆ 2.3.6 Before such high-risk AI systems can be deployed in the EU market, they will need to be subjected to stringent requirements providing for:

- adequate risk assessment and mitigation systems;
- high quality of datasets to minimise risks and discriminatory outcomes;
- the logging of activity to ensure traceability of outcomes;
- detailed documentation containing all information relating to the AI system and its purpose, to enable the authorities to evaluate its compliance;
- clear and adequate user information;
- appropriate human oversight measures to minimise risk; and
- a high level of robustness, security, and accuracy.[62]

◆ 2.3.7 In addition, before high-risk AI systems are put on the EU market, they will need to undergo approved conformity assessments and continue to comply with the requirements for such systems under the AIA. For certain AI systems, an external body will need to be involved in the conformity assessment audit. Where there are changes to the high-risk AI system, the conformity assessment process will be repeated. Standalone high-risk AI systems will be required to be registered on a designated EU database. Finally, a declaration of conformity will need to be signed and the high-risk AI system will have to bear a CE (*Conformité Européenne*) marking. And, after high-risk AI systems are put into the market, their providers must have an ongoing monitoring system to identify and address risks as they arise.

◆ 2.3.8 These requirements affecting a range of AI systems to be deployed in EU workplaces (many or most of which as currently defined are high-risk) will therefore be relevant to developers, providers and distributors of, and service providers and employers themselves using, such AI systems affecting employees and workers in the EU.

◆ **2.3.9 Limited risk:** AI systems that are subject to specific transparency requirements. The example of such requirements given in the Commission's own summary is: "[w]hen using AI systems such as chatbots, users should be aware that they are interacting with a machine so they can take an informed decision to continue or

[60] See the text of the proposed AIA, as cited in footnote 58.
[61] See the text of the proposed AIA, as cited in footnote 58.
[62] https://ec.europa.eu/commission/presscorner/api/files/document/print/en/ip_21_1682/IP_21_1682_EN.pdf.

step back".[63] Again, it is clear that some AI systems used in employment processes (as illustrated elsewhere in this paper) will fall into the limited risk category and be subject to these transparency requirements.

- **2.3.10 Minimal risk:** according to the European Commission, most AI systems will fall into this category. They will not be restricted under the AIA.
- 2.3.11 The AIA proposes very substantial penalties for the most serious infringements of its requirements – up to €30 million or 6% of global turnover. There are a range of other remedies and penalties proposed.
- 2.3.12 Although called the "Artificial Intelligence Act" the AIA will, like GDPR, have legal effect as a Regulation, laying down harmonised and uniform rules throughout the EU. And, as mentioned in paragraph 2.1.8 above, the AIA will also have extra-territorial effect[64], as it will apply not only to businesses and employers operating in the EU, but also to, among others: (i) providers who supply commercially, or put into service, AI systems in the EU, whether the providers are located in or outside the EU: and (ii) providers and users located outside the EU, if the output produced by the AI system is used in the EU. It is also worth noting that there are technical definitions of "providers" and "users" for the purposes of the AIA, with a hierarchy of legal and regulatory obligations starting with providers, and moving through to importers, distributors and users of AI systems.
- 2.3.13 The AIA is expected to be enacted as law by 2023/2024.

2.4. Examples of Current Laws and Regulations that we expect will apply to the Deployment of, and Reliance on, AI in the Workplace, but with Judicial Interpretation and Application

> "When we asked workers whether it was possible that the AI-powered technologies we had highlighted to them were being used at their workplace, but that they were just not aware of this, a shocking 89 per cent that responded said either "yes" or "not sure…We suspect this is largely due to a lack of consultation and transparency regarding the use of AI at work, and in relation to the collection, use and ownership of worker data."
>
> **- Technology managing people, The Worker experience, UK Trades Union Congress, November 2020, page 4[65]**

- 2.4.1 For greater clarity and to illustrate their potentially wider application beyond the jurisdictions concerned, in the following table we have arranged such examples thematically. Where practicable, they are drawn from several jurisdictions, also to make the point that the laws and regulations of many countries may apply in relation to the use of AI systems and processes in the workplace, and that employers need to consider as potentially applicable current laws and regulations wherever they operate and employ people. For this section, we have drawn from the thinking, taxonomy, and substantive contribution to this subject of Robin Allen QC and Dee Masters, the authors of *Technology Managing People – the legal implications*,[66] the value of which we wish to acknowledge here.

[63] As cited in footnote 58.
[64] As well as the "Brussels Effect": see footnote 40.
[65] https://www.tuc.org.uk/sites/default/files/2020-11/Technology_Managing_People_Report_2020_AW_Optimised.pdf
[66] https://ai-lawhub.com/technology-managing-people-the-legal-implications/

Theme	Representative examples of local laws and regulations, and the interplay with AI systems and processes used in the workplace
The legal basis of the employer-employee relationship	In the UK, under common law, the employment relationship is based, among other factors, on implicit mutual trust and confidence between employer and employee. Allen and Masters note that the main impact of this implied duty of trust and confidence is on the employer's obligations to employees. They point out that this mutual obligation of trust and confidence has significant consequences for the deployment of AI systems and processes in the workplace, including the following: • employers are generally obliged to provide explanations to employees for certain decisions taken by the employer; • the common law recognises that there is a power imbalance inherent in the employment relationship, the consequences of which include that the employer must take decisions about employees lawfully, rationally and in good faith; and • that, therefore, the conduct and actions of employers may often be subject to closer legal scrutiny than would apply to parties in an ordinary commercial contract. It follows that the deployment of AI systems and ADM in the workplace, especially where ADM impacts employees and their livelihood, will not release employers from their obligation to make decisions about employees that reflect the above requirements, both procedurally and substantively. Moreover, and apart from data protection laws like GDPR (referred to below in this table), employees may legitimately argue that certain decisions about them should be taken by a human being, not by an AI system or reliance on ADM, especially where human sentience and empathy may be required in decision-making. In any common law - or even civil law - jurisdiction where the same or similar legal principles govern the employment relationship, the above considerations may apply to varying degrees. This means that employers should satisfy themselves, and be able to demonstrate and explain to employees, works councils or unions, that the AI and ADM relied on by the employers satisfy the above requirements and principles, including that the AI systems and processes and ADM relied on have operated correctly, transparently, and can be adequately explained. In addition, in situations that could lead to the dismissal or disciplining of an employee, it would be prudent for the employer to provide for human review and an "override" of ADM.
Employee rights not to be unfairly dismissed	In many countries, the law protects some or all categories of employees from unfair dismissal. For example, in the UK there are statutory rights to fair process under Part X of the Employment Rights Act 1996, applicable to employees with over two years' continuous service. Such laws may, broadly put, be designed and applied to ensure both procedural and substantive fairness, e.g., to answer the following questions. • Did the employer follow all the prescribed steps in terminating the employment relationship? • In doing so, did the employer reach a fair, rational and lawful decision to dismiss the employee? • Can that decision be properly explained and therefore justified? Of course, as many illustrations elsewhere in this paper demonstrate, AI systems and processes do not, for various reasons, always function correctly and they may wrongly interpret data inputs for other reasons, such as bias. An employer's blind reliance on such systems and processes and ADM, together with the "black box" issue in AI, and the consequential lack of transparency and explainability in such decision making, may well render unlawful the employer's decision to dismiss an employee.

Theme	Representative examples of local laws and regulations, and the interplay with AI systems and processes used in the workplace
	Accordingly, employers' use of, and reliance on, AI systems and processes, including ADM, will need to comply, and be capable of demonstrating compliance with, fair process requirements and the need for correct, and therefore lawful, substantive outcomes in decision-making. It follows that employers: • will need to ensure through reliable audit processes that the AI and ADM have operated correctly and delivered, or have contributed correctly to, the right outcomes; and • should review the outputs and outcomes from those systems and processes and ADM and be ready to correct or override suspect outputs and outcomes, if necessary.
Equality laws, and discrimination in hiring processes, based on certain characteristics	Legal systems in various countries protect individuals and classes of employees and workers (including applicants for jobs) against various forms of discrimination based on certain "protected" characteristics, including (depending on the jurisdiction and legislation concerned) race, ethnicity, religion or belief, gender, sexual orientation, age, pregnancy and maternity, and disability. Some anti-discrimination laws distinguish between, and prohibit, direct and indirect discrimination (the latter including discrimination "by proxy", e.g., where a prospective employer may specify that a role must be performed full time, which could in effect discriminate against women who are often the main child carers, and therefore unable without great difficulty or at all to accept a full-time role). Where such laws apply, they may allow for some form of justification or defence, e.g., that the criteria applied are both job-related and represent a reasonable measure of job performance (as applicable under certain US anti-discrimination laws, as mentioned below), or that the employer or prospective employer had done everything reasonably possible in the circumstances to prevent the discrimination or disadvantage (as applicable under the relevant UK legislation). The consequences of proven unlawful discrimination in this context may include, depending on the jurisdiction and legislation concerned, unlimited compensation (as in the UK), fines, liquidated and punitive damages and reinstatement (as in the USA), and, more generally, legal fees and, of course, reputational damage, including the loss of customers and contracts and associated revenues through such damage. A breach of equality laws may therefore result in material adverse consequences for employers or prospective employers. We have explained elsewhere in this paper how bias in AI and underlying data sets arises, how AI systems and processes can and do result in detrimental outcomes, and how AI can discriminate against employees and workers. So there is no need to rehearse those explanations here. In the UK, the right to equal treatment and non-discrimination is a fundamental principle, which is enacted in the UK's Equality Act 2010 (with equivalent legislation applicable in Northern Ireland). In the USA, the principal laws are Title VII of the Civil Rights Act of 1964 (generally referred to as "Title VII"), the Age Discrimination in Employment Act (ADEA) and Americans with Disabilities Act (ADA, as amended in 2008). Title VII prohibits discrimination based on certain specified characteristics, e.g., race, colour, national origin, gender and religion. ADEA prohibits discrimination regarding age against anyone 40 years' of age or older. Both Title VII and ADEA prohibit discrimination based on disparate treatment or disparate impact, both of which have been held in US courts to include unconscious or implicit bias. Under ADA, among other provisions, it is forbidden in pre-hiring assessments to enquire about a candidate's physical disability, mental health, or clinical diagnosis.

Theme	Representative examples of local laws and regulations, and the interplay with AI systems and processes used in the workplace
	It has been suggested that: • a US court might decide that an employer relying on a biased algorithm that caused discrimination be held liable for that bias and the discriminatory outcome in its application, even if the bias actually reflected the bias of the AI's developer or programmer; • an employer could face a Title VII or ADEA disparate impact challenge if its reliance on the AI adversely affects members of a protected class. Apparently, such a challenge would be supported by *Griggs vs Duke Power Company*[67] and *Albemarle Paper Co. v. Moody*[68], in which the US Supreme Court held that, if standardised tests are shown to have a disparate impact on protected groups of employees, employers must establish that the tests are both job-related and represent a reasonable measure of job performance. So a US court might apply the same rationale to AI, compelling employers to demonstrate how the criteria analysed by the AI relate to the specific job requirements for the roles concerned; and • if AI used in the hiring or interviewing process discerns an applicant's physical disability, mental health, or clinical diagnosis, there could be a valid claim under ADA.[69] • We understand that the US EEOC has investigated situations arising from alleged AI bias in hiring and has stated that employers using AI systems and processes in hiring could face liability for any unintended discrimination. • In summary, in the context of discrimination and anti-discrimination laws and regulations, particular considerations in the deployment of AI systems and processes at the hiring stage and in related assessments include the following, some of which overlap to address different risks and concerns. • The need to identify, understand the impact(s) of, and to redress, bias inherent in algorithms and data sets. • That, aside from initial discriminatory effects that AI may create, it is likely to amplify those effects, which is why the first consideration above is critical. • The “black box” factor: the need for employers themselves to understand at a basic level of specificity how their hiring algorithms are working, including the factors being considered by the AI, which would in turn enable employers to know if the AI is likely to discriminate against individuals or groups. • Building on the last point, the need for sufficient transparency and explainability of AI in a hiring context, lack of which, from an employer’s perspective, might prevent it from relying on a defence or justification, if the employer did not itself understand how the AI had actually worked in specific situations, had not communicated to candidates its reliance on AI in the hiring processes, and/or was unable to explain how the AI had worked and therefore to justify its outputs and outcomes in specific cases. • Employers should monitor and test the outputs and outcomes of AI used in live hiring processes to observe and remove discriminatory operation and effects. • They should undertake appropriate prior due diligence of AI and data sets being used for hiring, including the competence, track record and reputation of AI developers, suppliers or service providers, and their voluntary commitment to ethical AI codes of practice.

[67] 401 U.S. 424 (1971).
[68] 422 U.S. 405 (1975).
[69] US legal and regulatory considerations in this part of the table and applicable situations are drawn from *Employment Law Red Flags in the Use of Artificial Intelligence in Hiring*, Gary D. Friedman, Thomas McCarthy, of the law firm Weil, published in *Business Law Today*, the American Bar Association, October 1, 2020: https://www.americanbar.org/groups/business_law/publications/blt/2020/10/ai-in-hiring/

Theme	Representative examples of local laws and regulations, and the interplay with AI systems and processes used in the workplace
	• Supply chain in AI systems and processes, and third-party involvement in the development and deployment of AI and data sets: employers should know who is responsible for the performance, outputs and outcomes of AI systems and processes at each stage of the hiring process, and determine whom to hold accountable for any resulting bias and/or discrimination and how any resulting liability can and should be allocated as between the employer and third-party suppliers or service providers • Employers' contracts with third-party developers, suppliers and service providers should reflect all applicable considerations outlined above.
Privacy rights and laws governing surveillance in the workplace	Under Article 8 of the European Convention on Human Rights (ECHR) and Article 8 of the UK Human Rights Act 1998, employees and workers[70] have a right to privacy. This right exists alongside and in addition to the data protection and privacy rights granted to data subjects under GDPR and UK's version of GDPR, the Data Protection Act 2018 (UK GDPR). Article 8 will apply to the surveillance and monitoring of employees and workers in the workplace and, as it seems likely, while they are working from home[71]. Accordingly, UK and other employees and workers having the benefit of Article 8's protection will have the right to a reasonable expectation of privacy at work, which will likely include their entitlement to private physical and virtual spaces. It is worth quoting Article 8 in full: i. *Everyone has the right to respect for his private and family life, his home and his correspondence.* ii. *There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.*[72] While as at the date of writing this paper, there appears to be no specific guidance on the application of Article 8 to the use of AI in the workplace, it will be interpreted dynamically in the light of changing standards in, for example, technological developments.[73] However, the extent of the application of Article 8 in relation to AI and related technologies to monitor employees and workers remains to be tested, along with the required justification presented by any employer deploying AI to legitimise their use of AI in this way. We understand that, in France and Germany[74], in addition to GDPR, there are legal and regulatory restrictions on surveillance by employers of their employees. In France, it appears that employees have a fundamental right to a private life, which encompasses the right to privacy, as well as for public, political and collective activities. And this right would extend to the confidentiality of certain communications. No system of surveillance or data collection may be installed without prior notice being given to employees

[70] Readers should note that "workers" (as opposed to "employees") enjoy certain protections under EU and UK law, though generally, in the employment context, such protection may not go as far as the protection afforded to "employees".

[71] According to the opinion of Allen and Masters, following *Barbulescu v Romania* (App No.61496/02, 2 September 2017 (GC): https://hudoc.echr.coe.int/spa#{%22itemid%22:[%22001-177082%22]}

[72] https://www.legislation.gov.uk/ukpga/1998/42/schedule/1

[73] https://ai-lawhub.com/technology-managing-people-the-legal-implications/

[74] For the examples under French and German law, we have drawn from *Artificial Intelligence is Watching you at Work. Digital Surveillance, Employee Monitoring and Regulatory Issues in the EU Context*, A. Aloisi, and E Gramano, published in Comparative Labor Law & Policy Journal, Vol.41: XXX, p101, first published October 25, 2019: Access | Comparative Labor Law & Policy Journal (illinois.edu)

Theme	Representative examples of local laws and regulations, and the interplay with AI systems and processes used in the workplace
	and to employees' representatives.[75] In French employment law (under the French Civil Code), there are apparently detailed rules that would govern the use of AI for surveillance and collection of employee data, including that employees must be informed about surveillance tools before they are deployed. Moreover, it appears that, under French employment law, employers must consult works councils about the introduction of new technology within the employer's organisation if this might affect employees' working conditions, employment, pay, training, and qualifications. In Germany, we understand that, in addition to GDPR, there is an overarching general civil right of personality under the German constitution, as shaped by German case law[76]. Certain German federal constitutional cases have emphasised the openness of the general right of personality to legal development, making this a dynamic right, which may well extend to any interference with, or curtailment of, that right through the deployment of AI and other technologies in the workplace. By virtue of the German right of personality, it appears that there is a fundamental right to the confidentiality and integrity of IT systems, which would include email. Data generated, processed, and stored by IT systems are to remain confidential, and cannot be covertly accessed or manipulated. Any covert surveillance measures must be limited to the protection of sufficiently material legal interests, where any threat to such interests is actually foreseeable. And there is also a principle of proportionality applicable to such surveillance measures, which would require them to be justified as proportionate in the circumstances.
Data protection	The EU's GDPR is recognised as the most prevalent data protection and privacy law globally. This is based on its uniform application throughout the European Union and the European Economic Area (EEA), and because of its extraterritorial effect[77]. Among other countries, Canada has adopted a similar law (the Personal Information Protection and Electronic Documents Act (PIPEDA)). In 2021 China's Personal Information Protection Law (PIPL) came into force. India is in the process of adopting a data protection law. Both PIPL and the proposed Indian data protection legislation have similar characteristics to, or are to an extent modelled on, GDPR. A detailed survey of GDPR is beyond the scope of this paper, but employers subject to the Regulation should have in mind the following, non-exhaustive, requirements in GDPR's application to the use of AI systems and processes in the workplace. These will apply to AI deployment at every stage of human capital management. • It is most likely that the introduction to the workplace of a new technology like AI and associated processes will give rise to a high risk to employees and workers, which would require employers to undertake a data protection impact assessment (DPIA) to enable them to identify and minimise the data protection risks of introducing AI. • AI processing must comply with the prescribed grounds for the lawful processing of personal data. There are special and more stringent rules applicable to special categories of data (e.g., relating to health, race and sexual orientation), which are clearly very likely to

[75] Code du Travail, art. L 1222-4 (Fr.)

[76] Aloisi and Gramano cite, among other authorities for this and related statements, a decision of the German Federal Constitutional Court (BVerfG), the "Elfes-decision" of 16 January 1957, 1 BvR 253/56 in their article referred to in footnote 74, in their footnote 99.

[77] Under Article 3 of GDPR, where data processing activities relate to the provision of goods or services to data subjects in the EU, or involve the monitoring of data subjects' behaviour as far as that behaviour takes place within the EU, controllers or processors outside the EU are subject to GDPR. And see paragraph 2.1.8 of this section and footnote 40.

Theme	Representative examples of local laws and regulations, and the interplay with AI systems and processes used in the workplace
	apply to the introduction of AI in the workplace as it might integrate with HR applications and employee and worker-related databases. • The AI must process personal data accurately and the data must be kept up to date. • Data processing by the AI must be minimised, i.e. limited to what is adequate, relevant and necessary for the purposes for which the AI is being deployed. • Data processed by the AI must be stored for no longer than is necessary for the purposes for which the personal data is being processed. • Processing of personal data using AI must be fair and transparent, which would require employers to give employees and workers specified information relating to the processing of their personal data by the AI, e.g. the kinds of data being processed, the purposes of the data processing, and the recipients of the data. • There must be appropriate security and integrity of the personal data being processed by the AI, including protection against unauthorised access and accidental loss and destruction of, or damage to, the data. Most of the above requirements will require, among other features, a high level of transparency and explainability of the AI used in the workplace, together with monitoring and auditing of outputs and outcomes. Employers will also need to have regard to compliance with GDPR where they have outsourced, or contracted out, the deployment and management of AI systems and processes. This consideration will also apply to supply chains, and sub-processors of personal data. A particular feature of GDPR (and its UK equivalent) that is of direct concern to the use of AI and ADM in the workplace are its requirements relating to automated decision-making processes under Article 22. Employees and workers have the right not to be subject to decisions "based solely on automated processing, including profiling, which produces legal effects". "Profiling" is defined as "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."[78] The application of the profiling provisions of GDPR to the use of AI and ADM in the workplace is clear. It is noteworthy that, to our knowledge, the first case on the application of Article 22 of GDPR arose from the workplace, when several Uber and Ola drivers, as gig-economy "workers", contested action taken against them by Uber and Ola through the deployment of AI and, apparently, ADM.[79] Under Article 22, employers are required to provide to their employees and workers, as data subjects, information about "the existence of automated decision-making, including profiling" and, in the case of solely automated decision-making under that Article, "meaningful information about the logic involved, as well as the envisaged consequences of such processing for the data subject".[80] Again, the application of these requirements to AI and ADM in the workplace is obvious.

[78] GDPR, Article 4(4).
[79] For the Uber drivers' case, see footnote 44.
[80] By virtue of GDPR Articles 13(2)(f), 14(2)(g) and 15(1)(h).

Theme	Representative examples of local laws and regulations, and the interplay with AI systems and processes used in the workplace
	There are exceptions to the right of employees and workers not to be subject to solely automated decision-making, which are where the decision: • is necessary for the entry into or performance of a contract; • is authorised by domestic law applicable to the data controller; or • is based on the individual's explicit consent[81] Regarding "consent", the UK ICO's guidance is that it must be "freely given, specific, informed and [an] unambiguous affirmative indication of the individual's wishes". And to be "explicit", "the individual should expressly confirm their consent, for example by a written statement, filling in an electronic form or sending an email."[82] The UK ICO makes the point that, for a data subject's consent to be specific and informed, the controller (in this context, the employer) must make a consent request explaining that the decision will be entirely automated. And, in its more detailed guidance on consent, the UK ICO states that "employers will need to take extra care to show that consent is freely given and should avoid over-reliance on consent."[83] Implicit in an employment relationship, given the imbalance of power in it, will be a concern that an employee's consent may seldom be entirely freely given. If the ADM falls under Article 22, employers are also required to adopt clear and simple ways for employees and workers to request, as is their right, human intervention in ADM or to challenge an automated decision, which will have to be implemented. And employers are also required to undertake regular checks to ensure that the AI systems and processes are working as intended. As mentioned in paragraph 2.1.10 of this chapter, the UK ICO, which is responsible for enforcing the UK's enactment of GDPR, will be updating its existing employment practices code specifically to take account of, among other issues, the increasing use of AI in the workplace.
Rights to associate and to collective bargaining and representation	Article 11 of the ECHR and the UK Human Rights Act 1998 grant every person the right to associate, including the right to join a trade union, and a right for unions to be heard and to pursue their members' interests, as well as for the unions' right to bargain collectively. Allen and Masters note that, in the USA, AI has been used, in effect, to undermine union activity by "screening out" job applicants who are likely to become trade union activists, identifying and predicting union activity at various work locations, as well as suppressing union-related online content.[84]
Health and safety at work	In the UK, relevant law and regulation includes the Health and Safety at Work Act 1974, the Working Time Regulations 1998, and other related subordinate legislation. Although this regime offers wide and detailed protection for employees' health and safety in the workplace, it appears that it has not yet been held to protect workers from the impact of AI systems and processes in connection with their mental health. This does not discount the possibility of such challenges being made, especially concerning the blurring of employees' personal and working lives, and the right not to be "always on".

[81] For useful and comprehensive guidance on UKGDPR and, by extension, GDPR, see the UK ICO's guidance: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/rights-related-to-automated-decision-making-including-profiling/.
[82] https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/when-can-we-carry-out-this-type-of-processing/#id1
[83] https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/consent/.
[84] As identified in the UK TUC's *Technology managing people, The Worker Experience*, last cited at footnote 65.

Findings on Workforce Surveillance from ShareAction Workforce Disclosure Initiative (WDI)

The Workforce Disclosure Initiative (WDI) aims to improve corporate transparency and accountability on workforce issues, provide companies and investors with comprehensive and comparable data and help increase the provision of good jobs worldwide. The WDI investor coalition is made up of 68 institutions, with US$10 trillion in assets under management.

In 2021, 173 companies responded to the survey.

On a newly added question this year, WDI shared its proprietary findings:

WDI indicator 2.12: Describe any workforce surveillance measures used to monitor workers, and how the company ensures this does not have a disproportionate impact on workers' right to privacy. If the company does not conduct any form of workforce surveillance, state this.

Summary

- The strong link between workforce surveillance disclosure and more extensive disclosure in general suggests that companies are progressively building towards gathering data on workforce surveillance as their workforce transparency improves.
- Companies provided relatively good levels of disclosure on workforce surveillance, despite this question being a new indicator to the WDI survey this year and a lack of requests for this data in other mandatory or voluntary sustainability disclosure frameworks and standards. This suggests that companies have already been gathering at least some data on the use of workforce surveillance internally, rather than providing this data as a result of well-established sustainability reporting requirements on this topic.
- The sectors and countries where workforce surveillance is most widely used, or not, are diverse, suggesting that surveillance practices are not necessarily tied to specific legal regimes or business models.
- Companies did not use the most intrusive forms of surveillance, such as home video surveillance and screen recording, potentially suggesting some level of consideration for the privacy impacts of these measures.
- Legislation, particularly the EU General Data Protection Regulation, was frequently cited as the basis for companies' approach to respecting workers' right to privacy, highlighting the importance of regulation in this area in shaping company practices.
- Too few companies are involving workers in their surveillance measures. For companies to be able to genuinely respect workers right to privacy, worker consent and involvement is crucial. Worker involvement in the design and implementation of surveillance measures can also help companies to ensure that surveillance measures are proportional to the intended purpose.

Detailed Findings

How transparent are companies about their use of workforce surveillance?

- 60 per cent of companies responded to WDI indicator 2.12: Describe any workforce surveillance measures used to monitor workers, and how the company ensures this does not have a disproportionate impact on workers' right to privacy. If the company does not conduct any form of workforce surveillance, state this. This is 17 per cent lower than the average for this section of the survey, Section 2: Risk assessment and due diligence, but just slightly below the overall average level of disclosure for the entire WDI survey, which is 68 per cent. This question is a new indicator to the WDI survey this year and data on workforce surveillance is not requested in other mandatory or voluntary sustainability disclosure frameworks and standards. Consequently, these relatively good levels of disclosure suggest that companies have already been gathering at least some data on the use of workforce surveillance internally, rather than providing this data as a result of well-established sustainability reporting requirements on this topic.

- There is a strong link between workforce surveillance disclosure and more extensive disclosure in general - companies that had been responding to the WDI for longer were much more likely to provide information for this question. 44 per cent of first-time responders to the WDI survey provided data, compared to 84 per cent of fifth-time responders, suggesting the companies progressively build towards gathering data on workforce surveillance as their workforce transparency improves.
- Transparency on workforce surveillance did not, however, correlate with actual use of workforce surveillance. Some of the sectors that had the highest response rate were those least likely to use surveillance, such as real estate companies, who had high levels of disclosure (75 per cent of companies responded to the question), but reported low levels of surveillance. This suggests that collecting this data is not predicated on actual use of surveillance practices and instead is linked to a more general awareness of workforce surveillance as a material issue for the business.

Who is conducting workforce surveillance?

- 56 per cent of companies reported using some form of surveillance, and there are significant disparities in the levels of workforce surveillance used across sectors. While no real estate companies and just 29 per cent health care companies conducted surveillance, 87 per cent of IT companies, 70 per cent of industrials companies and 67 per cent of financials companies did conduct surveillance. For some of these sectors, greater levels of surveillance are expected. For example, many financials companies reported the need to comply with specific kinds of security regulations as the reason for their surveillance practices.
- However, there were not necessarily clear unifying features between the companies with the highest or lowest levels of reported surveillance. There was diversity between manufacturing and service based companies in the sectors most likely to conduct surveillance (for example, both industrials and financials companies reported similar levels of surveillance). The same can be said when considering the country in which a company is based. The two countries where companies reported the lowest levels of surveillance were European (Germany and France, where 71 per cent and 54 per cent of companies did not conduct surveillance). In contrast, Spanish companies reported some of the highest levels of workforce surveillance, where just 40 per cent of companies did not report using workforce surveillance.

What forms of surveillance are companies using?

- Some forms of workforce surveillance are more popular than others. The most common forms of surveillance used were general digital surveillance (e.g., monitoring internet and email usage), which 17 per cent of companies that answered the question reported, followed by on-premises video surveillance (16 per cent of companies) and audio recording (e.g., of calls) (10 per cent of companies). The least commonly used forms were home video surveillance (e.g., through webcams), facial recognition software, screen recording, physical searches, and timing work, which no companies reported. These disparities may be due to the ease of implementing the more widely used measures and also the link between some forms of surveillance and legitimate business needs, such as quality management (such as recording audio of customer service phone calls) or logistics (monitoring where delivery vans are through geolocation). Companies did not use the most intrusive forms of surveillance, such as home video surveillance and screen recording, potentially suggesting some level of consideration for the privacy impacts of these measures.
- The kinds of surveillance used, unsurprisingly, corresponds to the nature of the business. IT companies were most likely to use general digital surveillance, such as email and internet use monitoring (38 per cent of IT companies reported they did so), which is unexpected given the digital nature of these businesses. Utilities and consumer staples companies were most likely to use on-premises video surveillance (25 per cent of companies from both

sectors reported they did so), which may be driven by safety considerations on potentially higher-risk sites such as warehouses. Utilities companies were most likely to use audio recordings of calls (38 per cent of companies) and geolocation tracking (38 per cent of companies), which many companies cited as being linked to the need to provide customer service and having fleets of vehicles regularly in use.

How are companies considering the right to privacy, data protection and worker involvement in their surveillance practices?

- The sectors that reported the highest use of surveillance measures were also more likely to provide data on measures to protect workers right to privacy, such as industrials and financials companies, which is encouraging. Legislation, particularly the EU General Data Protection Regulation, was frequently cited as the basis for companies' approach to respecting workers' right to privacy, highlighting the importance of regulation in this area in shaping company practices.
- Companies that conducted more surveillance also sought more worker engagement and input in data collection processes. However, levels of worker engagement were considerably lower than general considerations around data protections and workers' right to privacy, with just 11 per cent of companies providing data on this. For companies to be able to genuinely respect workers right to privacy, worker involvement is crucial. Worker consent to surveillance measures is vital in meaningfully respecting the right to privacy. Worker involvement in the design and implementation of surveillance measures can also help bolster this, providing companies with the insights required to ensure surveillance is proportional.

Chapter 3. Hiring

40% of Fortune 500 companies acknowledged that they are using some form of artificial intelligence in recruitment[85]. In October 2020, a group of ten senators, including Elizabeth Warren, has written a public letter to the Chair of the U.S Equal Employment Opportunity Commission (EEOC) to highlight their concerns over the use of employment technologies[86].

Specifically, "*effective oversight of hiring technologies requires proactively investigating and auditing their effects on protected classes, enforcing against discriminatory hiring assessments or processes, and providing guidance for employers on designing and auditing equitable hiring processes.*" This is not the first time US senators have shown concern about the issue. In 2018, another group of senators (including then-Senator, now the US Vice President, Kamala Harris) penned a similar letter to the EEOC, raising issues with the use of facial recognition technologies in the workplace[87]. Employment decisions in the United States are subject to the Uniform Guidelines created by the Office of Personnel Management (OPM). The Uniform Guidelines outlines the required validity

evidence which is admissible for showing that a selection procedure is appropriate and does not discriminate.

The EEOC is taking action. In September 2021, The Commission held a Zoom webinar[88] on the use of artificial intelligence and machine learning in employment decisions - a growing concern of federal regulators and lawmakers - while expressing optimism for AI's potential to reduce discriminatory outcomes. Commissioner Sonderling shared that when carefully designed and properly used, AI has the potential to advance diversity, equity and inclusion (DEI). Through broadened outreach channels, AI can open up opportunities to those who do not even know the opportunities exist; AI could therefore mitigate the risk of unintended discrimination. However, rapid and undisciplined implementation may perpetuate or accelerate such bias. This is because AI algorithms rely on a set of data inputs, such as resumes of high-performing existing employees, to guide job description and performance benchmarks for candidates. If those inputs lack diversity and skills for a future-fit business, AI algorithms may reinforce existing institutional bias. This is commonly referred to as disparate impact. Under the US jurisdiction, an employer is responsible for unintended discrimination and hence subject to penalty and disciplinary actions. On 10 November 2021, a bill[89] was passed in New York City to ban employers from using automated hiring tools unless a yearly bias audit can show they will not discriminate based on an applicant's race or gender: see paragraph 2.2.2 in the Legal and Regulatory Considerations chapter: *A Local Law to amend the administrative code of the city of New York, in relation to automated employment decision tools, 2021*.

3.1. Workforce planning

Workforce planning is a core business process which aligns changing organisation needs with people strategy. In traditional board director skills assessment, legal and audit expertise is ranked alongside business and executive management experience. Over the past ten years, technology (hardware and software), digital and cybersecurity expertise have become an integral part of the board skills matrix. More recently, sector specific sustainability expertise has been added to the mix. In particular, climate science expertise is most sought after.

[85] https://podcasts.apple.com/us/podcast/hired-by-an-algorithm/id1523584878?i=1000526571833
[86] https://www.vanhollen.senate.gov/imo/media/doc/Letter%20to%20EEOC%20to%20Clarify%20Authority%20to%20Investigate%20Bias%20in%20AI-Driven%20Hiring%20Technologies.pdf
[87] https://www.huntonlaborblog.com/2021/01/articles/eeoc-developments/senators-letter-to-eeoc-signals-scrutiny-of-ai-bias/
[88] https://eeoc.zoomgov.com/webinar/register/WN_MqDTHzZnQ6mA1hTOOwJsag?fbclid=IwAR0b8WCw_dKnrOGCY5EwLfweNWmQBqfNZbCvqelwUc90N2RUGlCrFw805MM
[89] https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4344524&GUID=B051915D-A9AC-451E-81F8-6596032FA3F9

Similarly, when human capital management identifies skills gaps, future skills need and disruptive intervention necessary to take a company forward, backward looking skills attributes and data from existing top performing employees alone will fail to fulfill the objectives of workforce planning.

Case Study 1: Boston Consulting Group customises workforce planning

- Name: Boston Consulting Group
- Headquarters: Boston, Massachusetts, United States
- Number of employees: Approximately 21,000

BCG is an early advocator of people analytics[90]. For its own workforce planning, it uses a range of AI analytics. One of them is Pymetrics assessment. It is an online gamified assessment in which candidates have 25 minutes to play through a series of twelve mini-games that take 2 to 3 minutes each to complete. It is an adaptive test that adjusts the conditions of each game based on answers from candidates.

Pymetrics is an employee analytics firm that uses cognitive research as a foundation and creates games by laypersons that assess the soft skills of candidates. These skills include attention, focus, effort, generosity, fairness, learning, decision making, risk tolerance and emotion[91].

Clients of Pymetrics such as BCG provide a set of online materials to address the frequently asked questions (FAQ) around fairness, transparency and the risk of bias in using this assessment[92]. Following the assessment, the candidate receives a report that documents the outputs of the assessment. The process by which bias is removed from the algorithm is open-sourced on GitHub to encourage open source bias testing[93]. The outputs are subject to gender and ethnic pass rates assessment comparison, also known as disparate impact testing. The threshold used follows the US EEOC guidance of 4/5ths or the 80% rule[94].

Training and coaching websites, such as those for the Scholastic Assessment Test (SAT) practice tests, have sprung up to help candidates prepare for these tests. According to a job assessment experts video, large employers such as Unilever, PwC, Accenture, MasterCard and KPMG also use gamified assessments[95].

- Reflections - AI in Workforce Planning

  On **Accountability -** Pymetrics Terms of Service Chapter 8 includes a Class Action Waiver, which means that candidates who accessed the assessment cannot sue the company for discrimination. An employer client of AI people analytics is ultimately responsible for any disparate impact; this is supported by the US EEOC law that regulates employers, not vendors and suppliers[96]. Therefore, companies that use third party analytics must conduct thorough due diligence, with clear documentation of the process and commit to ongoing assessment that such analytics are fit for purpose.

  On **Grounding -** BCG can benefit from being more explicit about the actual research literature that form the foundation of the people analytics used, some of them are already available online[97], such as an independent audit by academics at Northeastern University[98]. The audit suggests larger scale, observational studies based on the longitudinal data collected to evaluate model performance. Besides back-testing, more explanation on bias adjustment would also help strengthen the grounding of the AI algorithms.

[90] https://image-src.bcg.com/Images/BCG-Unlocking-Change-Management-with-People-Analytics-June-2019_tcm9-221713.pdf
[91] https://www.graduatesfirst.com/pymetrics
[92] https://media-publications.bcg.com/pymetrics-BCG-FAQs.pdf
[93] https://github.com/pymetrics/audit-ai and https://qz.com/work/1742847/pymetrics-ceo-frida-polli-on-the-ai-solution-to-hiring-bias/
[94] Questions and Answers to Clarify and Provide a Common Interpretation of the Uniform Guidelines on Employee Selection Procedures | U.S. Equal Employment Opportunity Commission (eeoc.gov). See question 11.
[95] https://youtu.be/cv-kLAcSUPg
[96] https://www.pymetrics.ai/terms-of-service
[97] https://www.pymetrics.ai/science
[98] https://evijit.github.io/docs/pymetrics_audit_FAccT.pdf

On **Bias -** Recommendation systems are purpose-built to find and replicate patterns in behaviour, updating predictions as employers and candidates interact[99]. If the system notices that employers prefer candidates with certain skills that are more commonly nurtured in private instead of state schools, such as public speaking, it may well prioritise proxies for skills that are more commonly found in candidates from wealthier backgrounds, exacerbating social-economic division. When a company signs up as a client, its selected current employees take an online Pymetrics assessment that is meant to extract the preferred attributes. Pymetrics then identifies patterns among the existing top performers in various roles[100] and generate different models for the roles.

In identifying the set of incumbent candidates to train a model, companies should be aware of the potential bias that may already exist in their internal performance evaluation systems. This is usually caused by legacy preferences that favour certain ethnic or gender groups. If the executive management of a firm has a lower proportion of certain sub-groups, it should recognise that there is under-representation in the training data set. The training data should therefore be subject to de-biasing processes before customisation. The challenge lies in how granular we need to go when creating these representative sub-groups. The U.S. Securities and Exchange Commission (SEC) amendments in business disclosure rules to include human capital management is a good start, but the disclosures may reflect a US-centric approach to categorisation, even for global companies[101].

On **Fairness** - According to Professor Arvind Narayanan from Princeton University, there are 21 definitions of fairness[102]. For example, group fairness could be different from individual fairness, which is determined entirely by the personal experience of the sample size of one, yet still valid through the eyes of the candidate. Statistical bias is a mathematical issue different from societal bias. If a hiring system is not focusing on disparate impact, but instead, focused on a different type of fairness, this should be specified. This is because a focus on other types of fairness may contradict with disparate impact. It is advisable for an employer to understand what specific form of fairness they are optimising for and ensure that this is used in the customised AI people analytics.

3.2. Recruitment

Recruitment is a process that covers role creation, writing of job description, role advertising, resume screening, shortlisting, interviews (including aptitude tests, language skills test and case studies, where appropriate), psychometric and cognitive tests and corporate values alignment assessment. Psychometrics cover three broad categories: cognitive ability or intelligence, personality or temperament, and mental health or clinical diagnosis.

Personality or temperament tests have established history in employment, many of which are based on scientific findings or academic research programmes that have controlled for potential discriminatory impact[103]. It is unclear how they are used to screen out a candidate as conclusions from those tests are generally used to assess team diversity rather than to optimise skills that match a particular job.

For example, general cognitive ability (GCA) tests are commonly used in predicting employee outcomes. However, these tests have been shown to contain group differences in the scores, meaning the test results contain bias. When used as a data point in hiring, this will drive decisions that lead to less diverse outcomes. Through a simulated study, it was shown that optimisation algorithms could reduce group differences in GCA scores by re-weighting test items to optimise for diversity and validity[104]. This way, the use of employee data in AI can become more objective through optimisation.

Disabilities — whether physical or mental — have been determined to be "private" information that employers cannot inquire about at the pre-employment stage, under the 1990 Americans with Disabilities Act (ADA).

[99] Bogen M (2019) All the Ways Hiring Algorithms Can Introduce Bias. See: https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias
[100] Todd S (2019) People are terrible judges of talent. Can algorithms do better? See: https://qz.com/work/1742847/pymetrics-ceo-frida-polli-on-the-ai-solution-to-hiring-bias/
[101] https://www.sec.gov/news/press-release/2020-192
[102] https://shubhamjain0594.github.io/post/tlds-arvind-fairness-definitions/
[103] https://hbr.org/2019/04/the-legal-and-ethical-implications-of-using-ai-in-hiring
[104] Allred, C. M. (2019). Applying a Metaheuristic Algorithm to a Multi-Objective Optimization Problem Within Personnel Psychology.

THROUGH THE EYES OF THE CANDIDATE – Voice from Rory Muldowney, a Neurodiverse Candidate

"From my personal experience, I believe that the use of AI in assessing and screening disabled/neurodiverse candidates through a pre-recorded video interview - conducting an assessment using factors such as body language, time taken to respond, eye contact, etc. - is discriminatory. It disproportionately screens these candidates out of the recruitment process due to their conditions.

An example of this would be screening candidates on 'tone of voice'. An individual who had a hearing impairment and/or speech impediment, for example, would usually have a tone of voice different from the average person of a population. If AI is used to assess candidates in a pre-recorded video setting on tone of voice, this would be achieved through using an 'if statement' (i.e. only allowing candidates whose tone of voice met certain criteria to continue with the application process). Hence, as many of said candidate's voices fall outside the 'desired' range of voice used by the AI system due to their condition, if the AI has only used neurotypical datasets for training and testing, a disproportionate number of said candidates would not pass to the next stage of the recruitment process.

I contend that, currently, an 'ethical' use of AI in the recruitment process for individuals with disabilities/neurodiversity, namely pre-recorded video interviews, would be for AI to not be used. My view is further justified from my personal experience of applying as a neurodiverse candidate to Summer Internships in Banking, specially Asset Management, through the charity 'EmployAbility' in 2019; the charity assists large companies, from Google to Magic Circle Law Firms, in fair recruitment practices for disabled/neurodiverse candidates.

Moreover, I highlight my applications for two US-based global financial groups' internship programmes to strengthen my argument. The charity EmployAbility created a separate application portal on their website to apply for the respective schemes, which removed the banks standard testing screens, such as the AI assessed pre-recorded video interview. Instead, a human recruiter would pick from the group of candidates who applied on the EmployAbility portal to complete the video interview and individually assess them, removing AI from the process.

Therefore, even the firms using AI screens in pre-recorded video interviews as part of the recruitment process appreciate that AI discriminates against candidates with disabilities/neurodiversity. Otherwise, why would they remove AI from the recruitment process?

The limits to conventional recruitment practices are well documented. Interviewer affinity bias (those who seek to employ those who look like them, with similar background and interests), inconsistent outcomes and less emphasis on job fit and employee satisfaction, could lead to higher than expected switching costs, both for the employer and the employee[105]. Predictive tools seek to help the employer make an offer that the candidate is likely to accept[106]. Unilever uses a combination of AI assessments, such as Pymetrics and HireVue, and acknowledged that it has hired a more diverse cohort of entry level job seekers, and reduced recruitment time[107]. Let us examine how this might be possible.

- AI in Candidate Search

  Before candidates have even applied for the job, AI methods can be used by recruiters and hiring managers to search for candidates. Firstly, algorithms can be used to target potential candidates. Such algorithms are optimised for visibility in the sense that they use likelihood of clicking on the job for the lowest cost as the primary criteria[108]. In order to optimise for clicks, these advertisements may target candidates considered to be relevant or likely to be interested in the job to get a greater amount of candidate applications. Secondly, AI is used in search through search engine algorithms that determine which individuals appear when recruiters are actively searching for relevant candidates. Such algorithms typically use relevance and ability of candidates as the main criteria where relevance may be the degree of overlap of skills from a candidate and those required for the job, and ability

[105] https://workofthefuture.mit.edu/wp-content/uploads/2021/01/2021-Research-Brief-Polli-Kassir-Dolphin-Baker-Gabrieli.pdf
[106] https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias
[107] https://www.businessinsider.com/unilever-artificial-intelligence-hiring-process-2017-6?r=US&IR=T
[108] Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A., & Rieke, A. (2019). Discrimination through optimization: How Facebook's Ad delivery can lead to biased outcomes. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1-30.

may be candidate feedback ratings or test scores[109]. Recruiters can use these algorithmic recommendations in choosing which candidates they will reach out to for the next stage of recruitment.

- Reflections - AI in Search

The **opportunities** presented by AI search are:

1. When algorithms are repaired to reduce bias, there are two advantages:
   a. **Opening up the hiring pipeline** to candidates who might not otherwise see a job advertisement and potentially reaching more diverse groups of candidates. For example, identifying talent in underprivileged communities, such as the new collar programme[110] . IBM has a tradition of such programmes including this in Brooklyn; and
   b. **Fair ranking algorithms** on search engines can improve selection of underrepresented groups for interviews[111], allowing more diverse candidates at the early stages of recruitment.
2. **Increasing efficiency** in the hiring process through human-computer interactions. Algorithmic recommendations were found to increase the fill rate of technical job positions[112] through recruiters using these recommendations in their hiring decisions.

The **risks** of using AI in search are:

1. Skews in job advertisements can occur where some demographic groups are less likely to see the advertisements than other groups. This can occur due to two primary reasons:
   a. **Economic factors:** there is a monetary budget for all advertisements, and therefore priorities could be set to optimise the likelihood of success. This can even influence advertisements which are programmed to be demographically neutral, which occurred with an optimised Facebook advertisement - found to be shown to 20% more men than women due to the price premium for advertising to women[113].
   b. **Stereotypes** built into algorithms can influence how the content of the advertisement is shown to certain groups. For example, an advertisement for a job in the lumber industry was found to reach an audience that was 72% white and 90% male[114] because it was judged to be more relevant for those groups.
2. **Bias in candidate rankings** on search engines can impact which candidates get invited to interviews, introducing bias at a very early stage in the recruitment process. In a study assessing three resume search engine websites, women were ranked statistically lower than men on up to 14% of job titles[115]. This means that when recruiters search for candidates, they are shown a greater proportion of men in the top rankings.
3. Many of the algorithms used on search engine and advertising websites are **"black box algorithms**" meaning it is difficult for users to understand how they work and make decisions. This can make them difficult to assess for potential aspects of the algorithm which might drive bias.

- AI in Resume Screening

AI applications in resume screening often use an **expert system**, one of the four AI applications discussed in Definitions**.** Early applications in human capital management examines its use as a decision aid through the partnership between domain and knowledge experts[116]. Generically, expert systems can be used for interview

[109] Horton, J. J. (2017). The effects of algorithmic labor market recommendations: Evidence from a field experiment. *Journal of Labor Economics*, *35*(2), 345-385.
[110] https://www.cnbc.com/2019/10/03/why-ibm-is-using-ai-to-find-jobs-for-people-without-college-degree.html
[111] Sühr, T., Hilgard, S., & Lakkaraju, H. (2020). Does Fair Ranking Improve Minority Outcomes? Understanding the Interplay of Human and Algorithmic Biases in Online Hiring. arXiv preprint arXiv:2012.00423.
[112] Horton, J. J. (2017). The effects of algorithmic labor market recommendations: Evidence from a field experiment. *Journal of Labor Economics*, *35*(2), 345-385.
[113] Lambrecht, A., & Tucker, C. (2019). Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads. Management Science, 65(7), 2966-2981.
[114] Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A., & Rieke, A. (2019). Discrimination through optimization: How Facebook's Ad delivery can lead to biased outcomes. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1-30.
[115] Chen, L., Ma, R., Hannák, A., & Wilson, C. (2018, April). Investigating the impact of gender on rank in resume search engines. In Proceedings of the 2018 chi conference on human factors in computing systems (pp. 1-14).
[116] Hannon et al (1990) The Feasibility of Using Expert Systems in the Management of Human Resources. See: https://core.ac.uk/download/pdf/5129172.pdf

scheduling, candidate selection and diagnostics, but they require all relevant knowledge to be coded in advance through an 'if-then' scenario, and therefore inflexible if the skills that are required for the job have many optimal combinations, and the ranking of these skills are not definitive and hence **dependent on how well the candidate can deploy them rather than possess them**. For example, one rule in a knowledge base developed by an employer to screen resumes for a role requiring a university graduate with over ten years of experience might read: "If the highest education level is less than university and years of experience is less than ten, then reject the candidate." Another rule might read: "If professional qualification equals No, then reject the candidate."

- Reflections - AI in Resume Screening

The **opportunities** presented by AI screening process are:

1. **Clear skills and experience requirements** in the job description make keywords and taxonomy explicit so it presents equal opportunities to applicants;
2. **Diverse considerations** on equivalent levels of certifications and have them embedded in the expert systems for screening;
3. Identify opportunities to **improve representation of under-represented groups** by integrating taxonomy of certain under-representation in the expert systems for screening.
4. **Automate resume blinding** to remove factors which contain group differences so that aspects of the resume which are analysed will not perpetuate group biases in hiring.

The **risks** of this screening process are:

1. **Discriminating against candidates** who fail to use the standard 'keywords' and the relevant 'taxonomy' in their resume;
2. **Embedded bias** that benefits certain categories of candidates due to cultural and language alignment advantages. It has been found that candidates who are linguistically similar to hired employees, measured by the frequencies of certain words used, have a higher likelihood of being hired by about 33-53%[117]. For example, if a UK based firm creates an expert system that specifies requirements for A level grades or degree honours rather than GPA or International Baccalaureate exam scores, those who took a non-UK high school examination may be disadvantaged;
3. **Group differences** stemming from aspects of resume experiences. For example, hobbies and interests can be a signal of being upper-class such as sailing, polo, and classical music. Past work has found that men who have these hobbies listed on their resumes are more likely to be invited to interview[118]. Thus, it is important that algorithms assessing resumes do not include characteristics such as hobbies that tie to group differences;
4. **Lacking in flexibility** that takes into consideration the optimal combination of skills which depend on the depth of experience rather than a formulaic approach that scores a candidate.
5. **Lack of uptake.** People in charge of hiring decisions may still prefer human recommendations over algorithmic recommendations. In a study of recruiters who were given both algorithmic and human expert recommendations about a sample of candidates, they were more likely to base their decisions around the recommendations from the other humans rather than the algorithms[119].

[117] Stein, S. K., Goldberg, A., & Srivastava, S. B. (2018). Distinguishing Round from Square Pegs: Predicting Hiring Based on Pre-hire Language Use (No. repec: ecl: stabus: 3627).
[118] Rivera, L. A., & Tilcsik, A. (2016). Class advantage, commitment penalty: The gendered effect of social class signals in an elite labor market. *American Sociological Review*, *81*(6), 1097-1131.
[119] Oberst, U., De Quintana, M., Del Cerro, S., & Chamarro, A. (2020). Recruiters prefer expert recommendations over digital hiring algorithm: a choice-based conjoint study in a pre-employment screening scenario. Management Research Review.

ZipRecruiter, an online employment marketplace[120] that uses AI for resume screening provides recommendations on resume writing that addresses some of risks highlighted[121]:

- Use short, descriptive sentences with keywords that are easy for AI to detect - however, each resume may need to be tailored to the taxonomy of each business;
- Highlight certifications as they are easily quantifiable;
- Highlight skills explicitly sought in the job description.

Other technology solutions that support candidates in the recruitment process include:

- **Jobscan**[122] that tests resumes and scores how well a resume can be matched by AI for a particular job;
- **Vmock**[123] highlights inconsistent dates and spot gaps. VMock's algorithm claims to help users identify transferable skills as well;
- **Interviewstream**[124] supports AI interview practices and record the mock interviews for candidates;
- **Biginterview**[125] uses AI to coach users on their interview responses;
- **Talk Hiring**[126] includes 10-minute mock interviews consisting of five questions each.

Unintended consequences and negative externalities:

Note that as more paid for services are available for candidates, there is a risk of pricing out poor candidates or applicants from a disadvantaged background without affordable access to digital resources to prepare for such screening.

[120] https://www.ziprecruiter.com/
[121] https://www-technologyreview-com.cdn.ampproject.org/c/s/www.technologyreview.com/2021/08/04/1030509/job-search-how-write-resume-ai-artificial-intelligence/amp/
[122] https://www.jobscan.co/
[123] https://www.vmock.com
[124] https://interviewstream.com/
[125] https://biginterview.com/
[126] https://www.talkhiring.com/

EXPERT OPINION – On Gaming the System - by Dr Mark McDonald, HSBC Global Markets

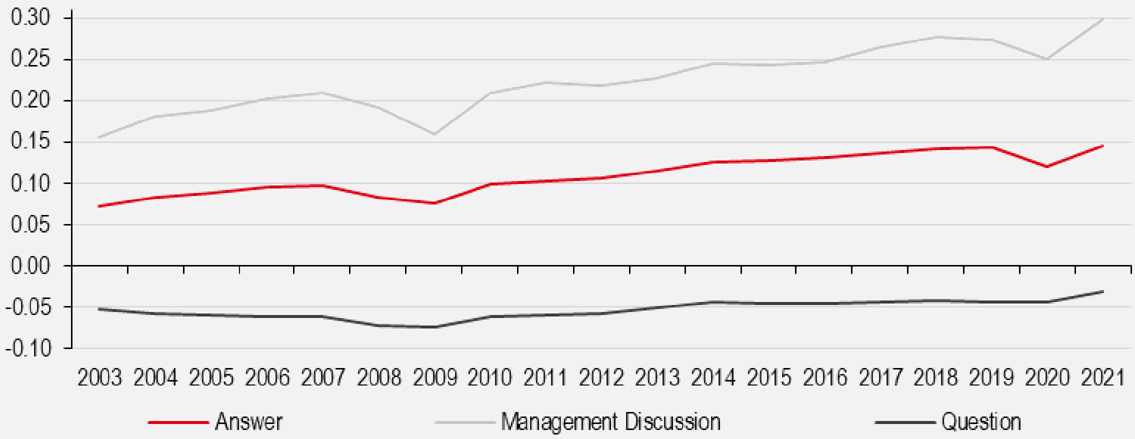
The risk of actors gaming an AI system with which they interact is not insignificant. When the consequences of the decisions made by an AI system are significant for an individual, it is likely that the individual will change aspects of their behaviour specifically to be scored more favourably by the AI system.

We can see this mechanism occur even with the most senior individuals within an organisation. When public companies release their results it is common for the company to hold a conference call which will be joined by analysts and investors. For many years now, the content of these calls has been transcribed by several data providers for the benefit of those who are unable to join the call as it happens.

In recent years it has become common for investors to analyse these earnings call transcripts using Natural Language Processing technology. A common piece of analysis performed using these transcripts is to assess the sentiment of company management. Since company managements are well aware of this practice, there is evidence that they are adapting their behaviour to combat these technologies.

Company earnings calls have three natural components: (i) a scripted introduction by company management (typically referred to as the "Management Discussion" section), (ii) questions by analysts and investors, and (iii) answers to these questions by company management.

Using analysis performed by the HSBC Global Research department, we show in the chart below the average sentiment of each section of these earnings calls through time. There is notable evidence of an upwards trend in this average sentiment from company management. This is most notable in the scripted Management Discussion, but is also clearly visible in the unscripted answers to questions. In contrast, the questions asked by analysts have shown negative sentiment on average.



Source: HSBC, Refinitiv TRKD

Changing behaviour of company management in the face of algorithmic assessment is also documented in Cao et al. 2020[127]. In that paper the authors show that companies whose corporate disclosures experience a high proportion of electronic downloads had significantly reduced the usage of words included on a widely used sentiment lexicon. However, the behaviour seen on earnings calls is more dramatic since it shows that company management have adjusted their behaviour even in unprepared answers to questions and not just in carefully scripted documents.

[127] (Cao et al. 2020) "How to talk when a machine is listening: Corporate disclosure in the age of AI", Cao, S.; Jiang, W.; Yang, B.; Zhang, A. L., NBER Working Paper 27950, available from https://www.nber.org/system/files/working_papers/w27950/w27950.pdf

◆ AI in Video Interviews

HireVue[128] provides video-based interviews that use algorithms to analyse speech content, tone of voice, emotional states, nonverbal behaviours, and temperamental clues. They are conducted with a computer camera or a smartphone. Candidates respond to a list of questions required for their role, speaking to the camera. HireVue's AI analyses the answers, noting aspects like keywords, body language, and tone. Hiring managers then see a detailed list of candidates the AI algorithm deemed performed best.

The **opportunities** presented by this interview process are:

1. **Efficiency:** Processing a large volume of candidates;
2. **Specificity:** Hiring manager can be explicit about the cognitive diversity and personality that they are seeking and use technology to support assessment.

The **risks** of this interview process are:

1. **Failure to comply with law:** as an example of such law, in 2.2.3 of the Legal and Regulatory Considerations chapter we outline the State of Illinois Artificial Intelligence Video Interview Act of 2019.
2. **Poor candidate experience if implemented badly:** Automated tracking and assessment is deemed damaging to mental health[129], especially if candidates are not given the details of how the models and algorithms categorised and analysed their behaviour. Many of the 'AI' assessment models are classification systems built on decades old statistical methods or classification model, such as support vector machine ("SVM"), as in the case of Pymetrics in Case Study 1. They are not built on what is commonly referred to as deep learning that involves layers of neural networks, therefore, there is a real opportunity for companies to be more transparent in disclosing the assessment methodology and methods to candidates.
3. **Embedded bias:** Academics have questioned if emotions and characters can be standardised from facial expressions[130]. It is also unclear how different cultural background is taken into account when gestures and expressions are 'assessed'. For example, in India, shaking the head from left to right signals a 'yes' or an agreement, whilst much of the world's population may take it as a 'no' or negative response.
4. **Fairness:** Researchers have found evidence that introverts are less likely to emerge as leaders[131]. Similarly, in interviews, introverts may not speak up as much or need more encouragement to put forward their views. Whilst a human-conducted interview by experts could emotionally sense such preference and adjust accordingly, it is unclear if the current AI video assessment is adjusted for such subtle personality differences, especially some analytics may use the tone of voice and measure the time used by candidates in putting forward their views as input indicators of assertiveness[132]. There are also other personality types that may be disadvantaged at standardised interviews, for example, one who tends to be more analytical in their approach, and therefore require more time and data to provide a thoughtful response.
5. **Concerns over fitness for purpose:** Large clients like Unilever with 142,000 employees may be well-positioned for AI analytics companies to build custom hiring assessments based solely on the richness and diversity of their proprietary employee data. Smaller clients may turn to pre-built assessments, making the assumption that the candidate pool and job role on which the assessment was built is sufficiently similar to warrant generalising its conclusions[133].

[128] https://www.hirevue.com
[129] https://www.theguardian.com/technology/2021/nov/11/algorithmic-monitoring-mental-health-uk-employees
[130] https://nypost.com/2021/03/04/hbos-persona-how-myers-briggs-and-ai-are-being-misused/
[131] Spark and Stansmore (2018) The failure of introverts to emerge as leaders: The role of forecasted affect Personality and Individual Differences 121:84-88 January 2018. See: https://www.researchgate.net/publication/322173544_The_failure_of_introverts_to_emerge_as_leaders_The_role_of_forecasted_affect
[132] Data-centric AI: Real World Approaches on YouTube by Andrew Ng, DeepLearning AI, 11 August 2021. https://www.youtube.com/watch?app=desktop&v=Yqj7Kyjznh4
[133] Raghavan et al (2019) Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices. See: https://arxiv.org/pdf/1906.09208.pdf

EXPERT OPINION – On Technology Mediated Mental Health Treatments – by Professor Elias Aboujaoude MD, Stanford University

"The use of AI in the provision of psychotherapy remains novel and has not been subjected to the level of testing that other technology-mediated mental health treatments have. The result is that affect, expressions of anxiety and other non-verbal communication have not been adequately analysed in the psychological literature, complicating the ability to rely on them in the work interview context."

Some companies measure sales agents' productivity and income using the following metrics:

1. Agent channel Net Book Value (NBV)
2. NBV per agent per year
3. Activity[134] rate of agents
4. Agent income per month

To become an agent, candidates must go through an AI interview processes. Their responses, voice, gestures and facial expressions are analysed, recorded and scored based on the attributes of past successful agents. Video recordings of interviews accumulated over the years form the training data of its proprietary AI system. The characteristics and behaviours of successful agents are analysed, integrated with productivity scores, which form different agents 'profiles'. Top performing agents display a diverse set of characteristics, skills and learning preferences, and they are 'assigned' to different ideal type groups. The on-boarding and skills enhancement training programmes are designed around their assigned ideal agent types.

Once candidates begin to apply for a given position, the system screens candidates' resumes and pushes the most suitable ones to frontline managers who can request online assessments and AI-powered interviews.

Companies that use this Smart HR system are likely to want to measure the return on investment in dollar terms using indicators such as time saved to hire (from search onset to new-hire on-boarding date); the value of improved efficiency; and the estimated return on on-boarding improvement (ROOI).

- Reflections - AI in video interviews

AI system works well for language specific interviews. As business expands into other markets, where regional dialect or different languages are spoken, and the local workforce combine English and other dialects speaking words in the same sentence during interviews, the system needs to be retrained to deliver the expected outcomes of return on investment.

In general, task specific AI often requires to be retrained when applied in a different sector or market, as well as over time.

As investors and companies seek more diversity and inclusion in workforce, it would be helpful to analyse how the overall profiles of their top-performing agents' cohorts change over time. This ensures that homogenisation does not become bottlenecks that prevent career opportunities reaching those who do not fit the ideal types of the past.

Furthermore, it is important that when AI is used in hiring, the prediction of performance is updated in line with changes in top-performers. This goes back to the type of AI used, where static AI will not naturally calibrate to these changes. Dynamic AI which is constantly updating the predictive model will be able to consider changes in top-performance and shift the factors in the prediction of best hires.

Facial expressions should be used with caution in video interviews. In a recent systematic review, looking at whether you can reliably infer emotions from analysis of facial expressions, found that there is insufficient

[134] Activity rate of agents = annual total of monthly agents who issued policies / annual total of monthly agents on board.

evidence to support this claim[135]. Further research needs to be done on whether accuracy can be improved in facial expression identification. Although there may be similarities in how people show different emotions, we cannot confidently infer an individual's emotional state through their facial expressions. When used in candidate assessments, individual variability may show up when certain candidates have good "poker faces", show their expressions easily, or fake certain emotional states. It is also important to consider whether inferring a candidate's emotional state is actually *relevant* for the job, and whether this violates an individual's right to privacy.

3.3. On-boarding

On-boarding is an essential aspect of employee recruitment. Successful on-boarding means enabling the new colleague to 'hit the ground running', generating good first impression of the employer, and accelerating cultural integration (See Chapter 3: Culture). However, most on-boarding process as we know are stressful for new employees, especially if the IT, Human Resources platforms are not connected nor coordinated. Benefits selection and procurement and expenses processes could sit on different supplier platforms that adds administrative burden.

According to the Society for Human Resource Management ("SHRM"), there are four components of successful on boarding, commonly referred to as the Four Cs[136]:

1. **Compliance** is the lowest level and includes teaching employees basic legal and policy-related rules and regulations.
2. **Clarification** refers to ensuring that employees understand their new jobs and all related expectations.
3. **Culture** is a broad category that includes providing employees with a sense of organisational norms - both formal and informal.
4. **Connection** refers to the vital interpersonal relationships and information networks that new employees must establish.

Case Study 2: Citizens Bank Jamie™ Conversational AI Virtual Career Assistant

- Name: Citizens Financial Group
- Headquarters: Providence, Rhode Island, United States
- Number of employees: Approximately 18,000

Citizens Bank Jamie[137], a conversational AI virtual assistant customised by Citizens Bank and developed by Paradox[138], won the 2020 BAI Innovation Award[139]. for its role in talent acquisition strategy, with customised candidate experience. According to Paradox, based in Arizona, United States, also partners with companies such as McDonalds and launched the hiring programme McHire for franchise owners[140] and with CVS Health, which also covers orientation programmes that is part of the on-boarding process[141].

Citizens Bank believe that Jaimie has created efficiency throughout the recruitment process as well as servicing employees. During COVID-19 period in 2020, Jamie helped to answer many questions related to diversity, equity and inclusion (DEI) – from mental health issues to recruitment process changes due to lockdown.

[135] Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological science in the public interest*, *20*(1), 1-68.
[136] https://www.shrm.org/foundation/ourwork/initiatives/resources-from-past-initiatives/Documents/Onboarding%20New%20Employees.pdf
[137] https://olivia.paradox.ai/co/CitizensBank/OfficeHours
[138] https://www.paradox.ai/
[139] https://www.bai.org/banking-strategies/article-detail/bai-global-innovation-awards-case-study-jamie-virtual-career-assistant/
[140] https://signup.mchire.com/
[141] https://olivia.paradox.ai/calendar/event/gAAAAABhVfnjKGofapbk_4U-S106zlHxfO1cho0z_HTCTY-Tqdo5LVxvZpPPvCFylhD38AnRRD0BDqfQQ0jdzbsfNv1_UQDOXw

Why is Jamie unique?

Jamie Chat uses natural language processing (NLP), robotics automation and features video responses, providing guidance on frequently asked questions such as benefits and corporate culture. As the recruitment and on-boarding interactions are captured and timely satisfaction surveys are provided, Citizens used it to measure satisfaction rates and set performance targets.

Jamie has a job search function that covers both internal and external candidates, including passive candidates where if profiles match, job opportunities may get 'pushed' to the potential candidates. This is done through matching candidate preference, desired location, and match suitable positions within the company. Jamie's job matching abilities enhanced the click through rate.

Jamie Scheduler is a recruiter calendar management function. According to Citizens, Jamie enables a more efficient scheduling process with 40% reduction in 'no shows' appointments and used Jamie to ensure that candidate feedback is provided in good time. The company also launched an employee referral process using Jamie based on the data and analytics outcomes that make the referral process simpler and more efficient.

Event management is a relatively new function adopted by Citizens and used in career fair. Given that many minorities candidates, such as Hispanic candidates, prefer to interact through mobile, Citizens believe that Jamie has helped to improve DEI in recruitment.

Who owns the conversational content?

When candidates interact with Jamie, all content is recorded and owned by Citizens.

What could be challenging?

Educating colleagues that automation is not about putting jobs at risk but to simplify administrative processes and repeatable tasks to create more time for career counselling advice.

A team with relevant skills and experience would need to be involved in training and integrating Jamie into daily routines on how AI and humans interact in order to deliver a personal and comfortable experience.

Change management processes should be put in place for the integration in order to ensure that the new systems are workflows are appropriately utilised in order to generate the return on investments as intended.

There is also a delicate balance between functionality, performance, compliance control and data privacy.

- Reflections - AI in on-boarding

  The **opportunities** presented by this on-boarding process are:

  1. **Self-help to simplify compliance and clarification tasks:** AI can minimise HR-related compliance tasks by automating the delivery and receipt of necessary paperwork. For new employees, automated tools can provide relevant sections of company policies and deliver login information relevant to a given role. AI can track that a document has been read and capture electronic signatures as new hires finish the necessary steps, removing the need for to confirm this completion manually. An AI-enabled virtual assistant can use natural language process (NLP) and understanding (NLU) capabilities to simplify.
  2. **Accelerating Cultural integration and Connection:** Using the unique employee profile, AI can customise recommendations for employee resource group (ERG) participation and learning and development courses (see Chapter 4: Performance).

  The **risks** of this on-boarding process are:

1. **'Black box' profiling**: According to the report Technology managing people - The worker experience by Trades Union Congress (TUC)[142], profiling could be an 'opaque process and susceptible to discrimination, even though it may be responsible for important decisions about people', as this involves condensing the richness of what a human being could offer, professionally, to a few selected attributes. On the other hand, this has traditionally been practiced through characters assessment such as Myers Briggs[143], a widely accepted form of personality assessment used for team building in a corporate context. If AI-driven profiling is used, the foundation and reference framework used could partially address the risk of the model operating in a 'black box'.
2. **Authenticity of alternative data:** This applies if a company uses alternative or social media data when building a profile. Candidate and employee data mining is particularly controversial despite the use of information based on semi-public or social media data that the candidate discloses voluntarily. Is Facebook, LinkedIn and/or Instagram Like a real like? Or does it reflect individuals who are reciprocating likes as a currency to demonstrate goodwill for targeted individuals, such as their CEO? Would the use of candidate alternative data create an image of what the candidate would like to project rather than a real profile that represents their genuine interests and opinions?

[142] https://www.tuc.org.uk/research-analysis/reports/technology-managing-people-worker-experience
[143] Unconscious Personality Bias Keeps Women From Leadership -- And Costs Companies (themyersbriggs.com)

Chapter 4. Culture

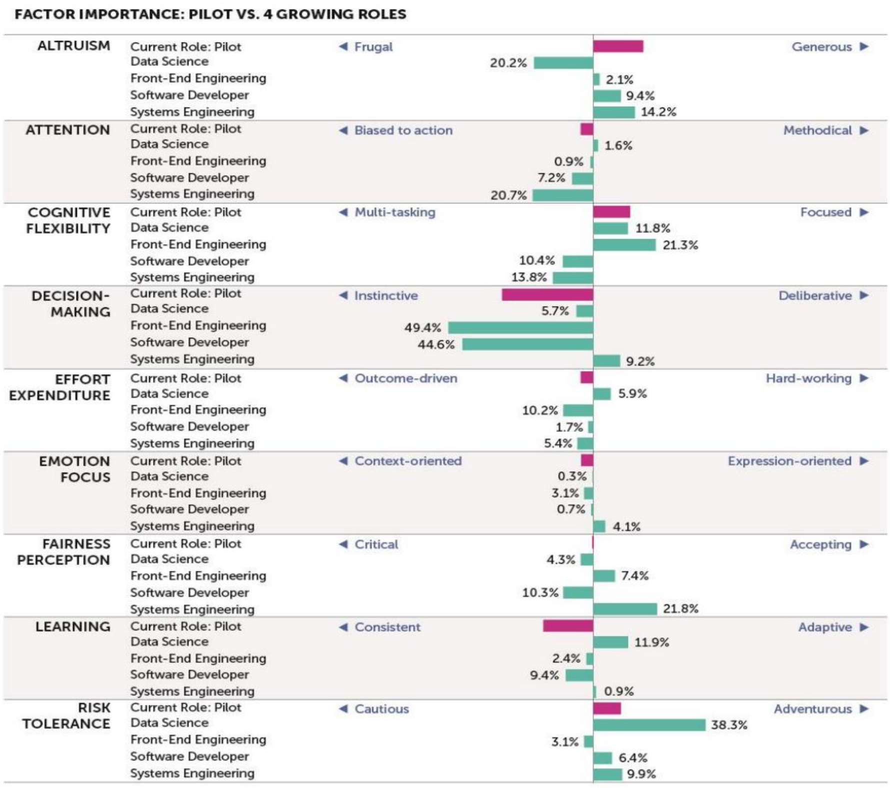
EXPERT OPINION – On Workplace Surveillance – by Professor Elias Aboujaoude MD, Stanford University

"Privacy mediates crucial psychological functions, such as resilience and the ability to bounce back from setbacks. Therefore, to minimise the risk of negative psychological impact from workplace surveillance, surveillance should be limited to the level absolutely necessary to measure performance and help with self-improvement"[144].

4.1. Competency

Competency management refers to systemically cataloguing, managing, and developing of job related skills. They are often role specific. Research suggests that soft and hard skills mapping encourage [mobility between different types of jobs][145]. This encourages agility amongst workforce, allow cross learning between departments, strengthening culture and collaboration across different businesses. Figure 5 outlines the soft skills of pilots and how they can be applied to growing roles in data science and software engineering.

Figure 5: Soft skills comparison between Pilot Role and Engineering Roles



[144] Aboujaoude E. Protecting privacy to protect mental health: the new ethical imperative. J Med Ethics. 2019 Sep;45(9):604-607. doi: 10.1136/medethics-2018-105313. Epub 2019 May 23. PMID: 31123190

[145] Polli F et al (2021) Cognitive Science as a New People Science for the Future of Work Research Brief 18, January 2021. See: https://workofthefuture.mit.edu/wp-content/uploads/2021/01/2021-Research-Brief-Polli-Kassir-Dolphin-Baker-Gabrieli.pdf

4.2. Team collaboration

Employees usually collaborate beyond their immediate departments and teams, especially for organisational projects and development of new products and solutions. When teams collaborate, there are risks of conflicts, such as different goals, priorities, approach and method of work. Identifying collaboration hotspots can help understand employee cohesion and improve organisational design, creating positive team dynamics.

The use of employee network analytics has increased during COVID-19 when team interactions moved from in-person to online experience where interactions can be tracked and monitored. For example, NetworkX[146] is a Python package that can be used to study the structure, dynamics, and functions of complex networks. It can visualise a network graph and perform analysis such as centrality, clustering, similarity measures, distance measures and link analysis (Figure 6).

Figure 6: Team Collaboration Metrics Framework

Metric	Description in Machine Learning	Higher Value Means
Clustering coefficient	Fraction of pairs of the node's neighbours that are adjacent to each other	Forming a close group e.g. early career employees from the same cohort
Centrality	Percentage of nodes connected with this node	Multiple direct connections e.g. A manager with their team members
Betweenness	Percentage of times this node acts as a bridge between groups	Coordination between groups e.g. project manager

AI can help to enable teams in working together more effectively through a couple of ways. First, AI can be used to predict which people will work together effectively. As Deloitte puts it, integrating aspects of automation and augmentation has the potential to create "superteams" which leverage human productivity[147].

In a study that looked at how AI can maximise collective intelligence (CI); the shared group intelligence that emerges through collaboration, three types of assisting AI were assessed;

1. A real-time feedback tool which quantifies the contribution made by each team member;
2. A to-do tool which guides the team on task allocation;
3. A chatbot to guide members on application of their skill and knowledge.

It was found that all types of AI interventions increased collective intelligence in the teams[148]. However, members also reported that the AI was at times unhelpful and distracting. Thus, it is important to keep user experience at the forefront of products aiming to assist teams.

Case Study 3: Nestle uses Swoop Analytics for Workplace from Meta

- Name: Nestle
- Headquarters: Vevey, Vaud, Switzerland
- Number of employees: Approximately 273,000

[146] https://github.com/networkx/networkx/issues
[147] https://www2.deloitte.com/us/en/insights/focus/human-capital-trends/2020/human-ai-collaboration.html
[148] Gupta, P., Kim, Y. J., Glikson, E., & Woolley, A. W. (2019). Digitally nudging team processes to enhance collective intelligence. In *Proceedings of the 7th ACM Collective Intelligence (CI) Conference. ACM.*

Nestle created a Connected Leaders programme on its Enterprise Social network (ESN) and uses SWOOP Analytics as a tool to measure team collaboration[149]. This programme was launched during the pandemic in a time of hybrid working when teams faced new challenges connecting and communicating digitally.

SWOOP Analytics is a provider of social networking and relationship-based analytics. It analyses the content and relationships on Microsoft Teams, Yammer and Workplace from Meta to provide organisations with collaboration and behavioural insights.

SWOOP uses a goal setting function[150] as part of their enterprise social networks which the organisation can alter to fit their needs. They base this off prior work identifying metrics of engagement which relate to having a business impact[151]. The metrics they use are as below, and can be provided at the individual, group, or enterprise level (Figure 7):

Figure 7: Team Collaboration Metrics Examples

SWOOP Metric	Description
Interactive users	Percentage of users who have taken a tangible action such as a post, reply or like. Reading, browsing, joining or leaving group is not included.
Response Rate	Percentage of posts that get a reply.
Public Messages	Percentage of messages posted in public groups.
Curiosity Index	Percentage of messages (posts and replies) that includes a question.
Mention Index	Percentage of posts that include one or more @ mentions.
Multi-Group Participation	Activity participation in multiple groups. The Diversity Index at the enterprise level is the average of scores for each person.
Two-Way Relationships	Calculates the proportion of connections that are reciprocated.
Influencer Risk	This measures how reliant the organisation is on a selected few power networkers.
Persona Distribution	The distribution of people by 'active' SWOOP Personas, i.e. Broadcaster, Responder, Catalyst and Engager. Observers are excluded.

Anyone in Nestle can reply to a Workplace post and provide an idea. Nestle creates digital leaders who work with the corporate communication coaches to refine their messages.

During COVID-19, Chris Johnson, Executive Vice President and CEO for Asia, Oceania and sub-Saharan Africa (AOA) created a post with video to thank frontline workers for their hard work and encouraged colleagues to share his message and add their own words of thanks.

In order to enhance engagement on posts and campaigns, SWOOP can identify the top influential people in an organisation who can be recruited to amplify messages. This in conjunction with SWOOP analytics on the best time to post can help to ensure that engagement is at maximum levels. In assessing team collaboration, SWOOP network analysis can spot gaps in communication at the team or organisational level.

4.3. Reflections - AI in Team collaboration

The **opportunities** presented by this team collaboration process are:

[149] https://www.swoopanalytics.com/case-studies/nestle
[150] https://www.swoopanalytics.com/blog/goalsetting
[151] Langen, M. (2015, August). Social collaboration metrics. In Companion to the Proceedings of the 11th International Symposium on Open Collaboration (pp. 1-4).

1. **Impromptu and timely communication:** the platform connects leaders to frontline workers as if he was speaking to them directly, and emotions and feedback from frontline workers are captured and analysed. There is increasing use of 'Pulse' surveys where employees are asked to share their moods and opinions in weekly quick surveys to measure the 'mood of the crowd'.
2. **Information for a digital working era:** In the hybrid and remote working world where we lack information previously acquired through in-person interactions, SWOOP analytics on engagement and network position can be rich sources of data regarding how teams are functioning. It can be used to target areas of improvement and spot individuals at risk of low levels of engagement.

The **risks** presented by this team collaboration process are:

1. **Supply chain transparency and human risks** - SWOOP employs 'Miners' to do the work, who are paid on hourly basis. There is a need to assess working hours, workplace practices and data privacy protocol of the vendor. There is a need to ensure that background checks have been conducted for these roles, any human rights due diligence assessments have been conducted especially around minimum and living wage compliance in the supply chain.
2. **Authenticity questioned**: Do we find that the key workplace influencers tend to be senior executives? Do most people respond and click on 'like' because they genuinely support the idea or was it just a way to earn 'brownie points'? AAAs Thomas Kuhn, an American physicist family said - 'The answers you get depend on the questions you ask.' Can pulse survey be gamed?
3. **Unknown impact and bias against certain behaviours**: Do extroverts and introverts behave differently when using online chats and workplace postings? Are they different one on one when compared to group chats?
4. **The impact of surveillance**: When employees are aware they are being surveilled by their organisation, they may act differently, rendering some of the analytical metrics obsolete. They may feel a lack of trust from the organisation towards them as a source of motivation for the surveillance.

4.4. Mobility

Mobility management refers to re-location within a company to take on new roles in a different geographical location, business entities or departments. It could also mean that by using transferable skills-based assessment, the pool of candidates can be expanded, improving diversity of background and skills developed in different industries.

The **opportunities** presented by AI in mobility are:

1. **Equalise access to opportunities** through employees having equal chances to hear about job opening. AI can recommend jobs that individuals might not have considered or heard about through word of mouth.
2. **Improve retention.** Individuals may be more likely to stay at a company where they can be matched into a role that aligns more closely with their future career aspirations.
3. **Enabling culture within an international organisation**: large global organisations have many sub-cultures – by geographies, business groups, or business functions. Increased mobility strengthens organisational culture through seconded individuals who have the internal network to connect different businesses, improve consistency, minimise silos and develop more comprehensive practices to manager inter-group conflicts of interests.

Chapter 5. Performance

5.1. Remuneration

Currently, most pay and promotion decisions are made by managers who work closely with the individuals they manage. However, unconscious biases can creep into these decisions when the impact of likeability or similarity can hinder objectivity in these decisions. Especially with the shift to remote work, it is imperative to ensure that pay and promotions are not based on visibility which might occur when certain groups of employees decide to work in the office, and others choose the flexibility of at home work.

AI can be used to assist in making complex compensation decisions that humans would not be able to without the help of algorithms. When a multitude of factors influence the pay of an employee, algorithms can be a tool to translate them into quantified outcomes. Furthermore, algorithms can take into account the supply & demand of certain employee skills based on labour market trends and compensate employees accordingly.

Can AI improve pay transparency around variable pay and promotion?

The **opportunities** presented by AI in performance management are:

1. **Improve employee perceptions of fairness through transparency.** Currently only 40% of employees believe their pay is fair[152]. Transparency through giving employees more details surrounding how their compensation is determined was found to increase ratings of fairness. When compensation decisions are based on objective criteria, such as those used in AI, employees can learn about the factors that directly relate to their pay.
2. **Remove pay and promotion gaps.** AI can detect gender and ethnicity pay gaps and be programmed to actively reduce gaps between any demographic groups. This involves basing pay and promotion decisions on factors which do not contain group differences.

The **risks** presented by AI in performance management are:

1. **Unable to capture the unique contributions of employees.** This is a fear that unique employee efforts may go undocumented if they are difficult to quantify. If employees feel this way, it is possible that a lack of morale may occur where employee efforts are only targeted towards factors the AI considers. An option to mediate this is to have human managers adjust the AI pay level based on factors that might not be considered in the algorithm.
2. **Historical pay inequities** being programmed into algorithm. Thus, it is important that AI pay assessment not only be based on past predictors of performance, but also the future in demand skills from workers.

5.2. Attrition

Employee attrition is a key concern for companies because too high an attrition rate could lead to loss of expertise, loss of productivity, customer relationships and increase hiring and training costs.

Data that is required to create an AI model for employee attrition prediction includes education and work history profile, tenure, performance and behaviour ratings, compensation, pay increment, promotion history and relationships of the employee with their manager and colleagues. Then a classification algorithm is applied, often built on open-source software library, such as TensorFlow.

[152] https://www.gartner.com/en/human-resources/trends/should-machines-make-pay-decisions

5.3. Learning

AI in Learning is often a 'self-help' function for individual employees. The most common technique used is user item recommendation, a collaborative filtering technique to identify similar users based on items that are common. We are familiar with this based on our online shopping and e-commerce experience under the recommendation titled:

'Frequently bought together';

'Compare with similar items';

'You might also like'; and

'What other items do customers buy after viewing this item?'.

Companies can use employee IDs, learning course codes and ratings as data, converting text to numbers, and use natural language processing (NLP) techniques to read comments, assess sentiment and discover relationships between user and learning modules. The outcomes can then be used to build a course rating prediction model for optimising suitable learning courses customised to each employee.

AI can match skills to roles and hence well-positioned to personalise learning programme. However, learning should be targeted at developing certain skills set instead of having a rigid list of skills desired for each role. We recommend that learning programmes should be skills based plus role based to ensure flexibility and transferability of skills are supported e.g. no role-based stereotypes, Company should demonstrate that learning recommendations are integrated with mobility performance and remuneration.

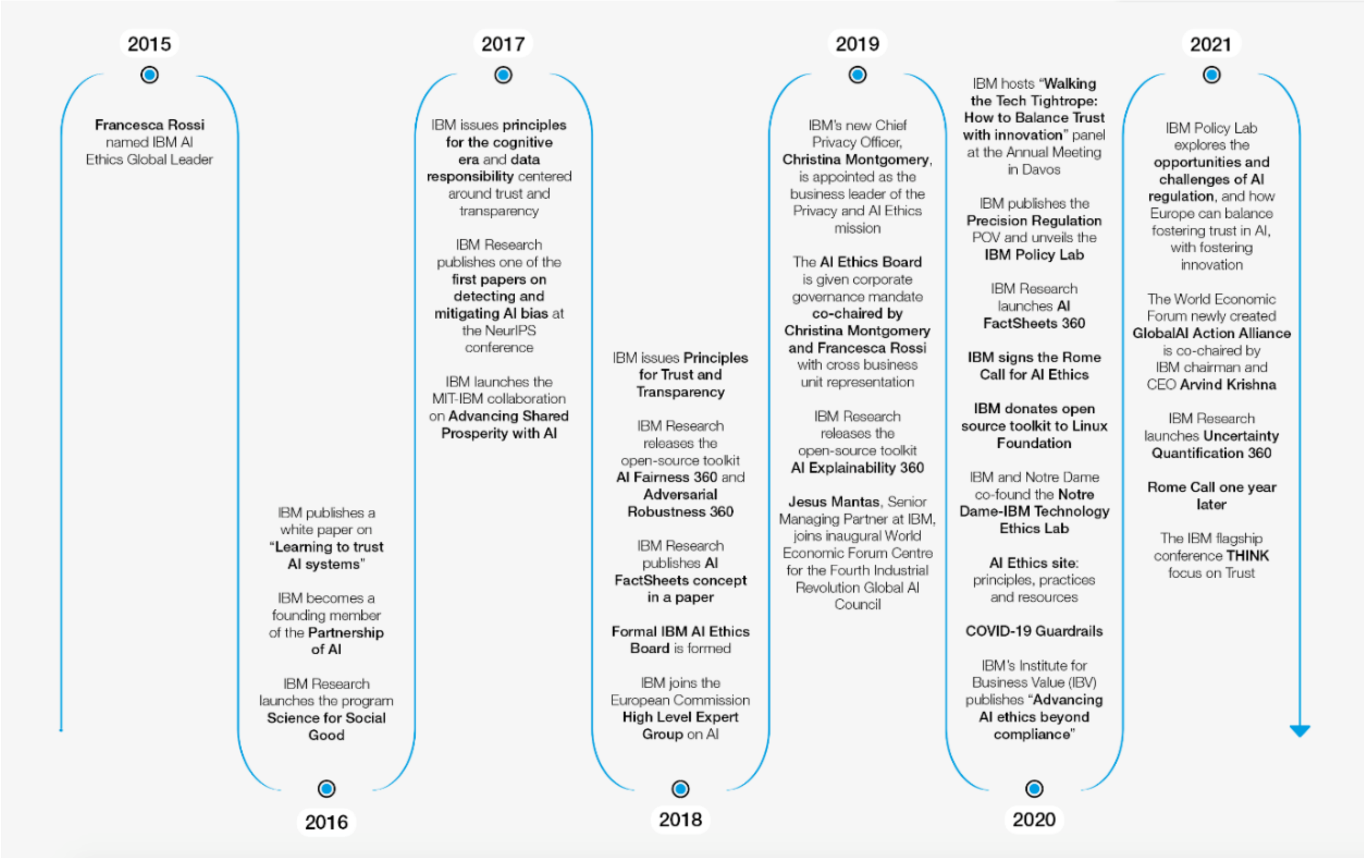
Case Study 4: IBM – Skills as the Silver Thread in the Talent Lifecycle

- Name: IBM
- Headquarters: Armonk, New York, United States
- Revenue FY 2020: US $73.6 Billion per 2020 annual report
- Number of employees: More than 250,000

The Evolution of Artificial Intelligence Ethics at IBM

IBM has made their position known on ethics and leads in workplace diversity, inclusion and equality, with a history of fair and inclusive practices including:

- 1911: Black and female employees included from IBM's founding.
- 1935: T.J. Watson, Sr. stated women will do the "same kind of work for equal pay" as policy.
- 1984: Non-discrimination on the basis of sexual orientation included in IBM's equal opportunity policy.
- 2021: Practicing radical transparency, IBM releases hiring, pay and promotion statistics for diverse populations at the company. At IBM, women earn $1 for every $1 earned by men for similar work. The same is true for underrepresented minorities in the US.



Source: IBM AI Ethics case study153

A Principled Approach

IBM's use of AI for the workforce is governed by its **Principles for Trust and Transparency:**

- The purpose of AI is to augment – and not replace – human intelligence.
- Data and insights belong to their creator.
- New technologies, including AI, must be transparent and explainable.

The following pillars are used to guide the design, development and use of AI:

- **Explainability:** AI system's ability to provide a human-interpretable explanation for its predictions and insights.
- **Fairness:** Properly calibrated, AI can assist humans in making fairer choices. An AI system should provide equitable treatment of individuals or groups of individuals.
- **Robustness:** As systems are employed to make crucial decisions, AI must be secure and robust.
- **Transparency:** Transparency reinforces trust, and the best way to promote transparency is through disclosure.
- **Privacy:** AI systems must prioritize and safeguard consumers' privacy and data rights.

These principles and pillars together demonstrate IBM's longstanding commitment to take an ethically conscious approach when bringing new technologies into the world.

[153] WEF_Responsible_Use_of_Technology_The_IBM_Case_Study_2021.pdf (weforum.org)

IBM's Skills-based People Strategy

Competitive skills sets are evolving rapidly in today's environment. To keep up in the workplace, employees and candidates expect consumer-grade experiences and equitable access to opportunity. These consumer-grade experiences require personalisation, making AI not a luxury, but a business imperative.

IBM has built AI solutions for HR designed to foster trustworthiness and deliver an end-to-end employee experience across the entire employee lifecycle. It helps to redefine how it attracts, develops and retains talent to transform the human experience.

Solutions are designed to deliver personalised experiences, with skills as the silver thread across the employee journey, empowering IBMers to make fact-based decisions. Key experiences are explainable, they empower users to give feedback, and they help identify and mitigate bias.

IBM's tools help managers build a skilled, diverse and inclusive workforce. Reflected within are IBM's ethical principles and those of trustworthy AI154: that AI augments human intelligence without replacing human decision making; that data and insights belong to their creator; and that AI systems must be transparent and explainable.

- **Sourcing candidates:** AI augments recruiters and hiring managers by actively guiding them through the process and ensuring that diverse pools of candidates are being considered for job openings.
- **Assessment suite:** Assessments augment, but do not replace, the hiring manager's decisions. For example, a game-based assessment measures fluid reasoning, which encompasses the candidate's ability to solve new problems without prior knowledge. This ability influences the capacity to learn quickly on the job and succeed in a role. The process is designed with human oversight, with the recruiters ultimately accountable for assigning candidates to the slate.
- **Interviewing:** Hiring teams receive intensive guidance on how to avoid bias when selecting candidates.
- **Selection:** The system guides managers to incorporate diverse perspectives in the interview process to help evaluate skills and sends hiring managers further best-practice reminders.

The result: Quality of hires rose 10% year-over-year in 2020, and the hiring of US underrepresented minorities rose 20% over the last three years[155].

Learning, Career Development & Opportunity

IBMs integrated digital career experience helps IBMers reflect on their skills, take learning that helps them develop, grow mentoring relationships and find jobs that help them meet their career aspirations. This experience exemplifies IBM's 5 pillars of Trustworthy AI:

- **Transparency:** IBMers can reflect on their skills using skills inference, that is based on a machine learning algorithm using data from IBMer's internal digital footprint such as resumes, learning, badges, research papers, etc. The result is shared back with the employee and the employee can provide feedback with over 1 million feedback points from employees. A consumer-grade experience similar to hailing a cab or watching movies on a streaming service provides personalised learning recommendations which employees can choose to take.
- **Explainability:** IBMers are given the access to understand data sources and how the AI works in their internal digital experience. They receive personalized learning and job recommendations with explanations so users can understand context, for example. the match can be based on alignment to the employee's current role or their next desired role.
- **Fairness**: IBM has a repeatable approach to identify potential biases and mitigation. This evaluation is conducted anytime the model changes.

[154] https://www.ibm.com/watson/trustworthy-ai
[155] https://www.ibm.com/impact/be-equal/pdf/IBM_Diversity_Inclusion_Report_2020.pdf

- **Robustness**: Operational rigor drives the system as it handles data abnormalities and exception conditions enabling confidence in the outcomes.
- **Privacy**: Employee data is sacrosanct. IBM follows all country-specific regulations such as GDPR with systemic approaches and education to ensure awareness.

The result: 98% of employees used the IBM learning platform every quarter and has a Net Promotor score of 58.

Remuneration

AI provides recommendations that managers can choose to use to help them make objective skill-based pay decisions and foster transparent pay conversations aligned with trustworthy AI's pillars of governed data and AI technology: transparency, explainability, fairness, robustness, and privacy.

Ultimately, managers know their employees best, so while AI provides recommendations, it does not automate the decision. The manager is the decision-maker. Managers view the explanations for these recommendations and may use them to help inform and shape the conversation with the employee.

- **Transparency:** Employees have access to their pay relative to the market. This transparency fosters evidence-based conversations between the employee and manager. IBM is creating AI FactSheets which, like food nutrition labels, provide a framework to document machine learning models and AI services and discuss how the model was created, tested, deployed and evaluated; how it should operate; and how it should, and should not, be used.
- **Explainability:** The system provides managers with salary increase recommendations tailored for each of their employees – each has the reasons supporting and explaining the individual recommendation for managers' consideration. For example, attrition risk using factors such the employee's job role, pay relative to the market, and historical attrition rates for the country or business unit, helps managers understand attrition patterns.
- **Fairness:** IBM uses AI Fairness 360, part of IBM's open-source toolkit, to help examine the machine learning model for potential bias identification and mitigation.
- **Robustness:** In addition to upholding data privacy commitments, IBM upholds operational rigor around design and use of AI for pay recommendations including the handling of data abnormalities and exceptions. IBM has created and deployed a foundational training module for HR Professionals that includes outlines what it does and what it does not do when building and training models.
- **Privacy:** Employee data is sacrosanct. Rigorous processes and accountability help ensure that the use cases covered by all projects adhere to employee and data privacy requirements.

The result: Attrition was reduced by one-third when managers followed the recommendation.

"It used to be that you'd learn a skill and it would last you for 30 years," says Nickle LaMoreaux, IBM Senior Vice President & Chief Human Resources Officer. "Now, skills in the tech world have about a three-year lifecycle." Showing employees the market demand for their skills can encourage them to acquire new ones, she says.

"It's really important to base pay decisions not only on the skills that you have, but on future skills and the supply and demand for those skills," says LaMoreaux. "That's a hard shift for managers to make because they're used to making these decisions based on performance. "

5.4. Reflections - AI in performance management

Personalisation and consumer grade experience at work is an attractive proposition for employees, and AI is especially suitable to identifying new patters, so that when applied correctly, it will expand the opportunity set for all employees.

When giving managers advice on pay increase based on skill set in demand and performance, it is important that the AI tool is only providing data as decision-support, and the ultimate responsibility rests with the manager.

The trustworthy AI journey is a marathon, not a sprint. We need to continue to build the expertise of practitioners and to push the envelope in the creating and fostering of trustworthy IBM AI FactSheets explain to end-users how each AI solution is created, tested, deployed, and evaluated. Together with other stakeholders in the ecosystem, it needs to build and adopt industry-wide standards for AI in HR, so we can all make progress together.

5.5. Conclusion

The potential of AI is enormous, as long as we are realistic about our expectations, and understand that it does not represent a magical wand; far from it, it requires humans to confront the deepest and darkest parts of human nature, such as our pride and prejudices, and be explicit about all possible scenarios of future and rank them in a way that is acceptable to the key stakeholders.

Clearly, this is almost an impossible task. However, what seems impossible does not mean that we shall shy away from the opportunities technology has to offer. On the contrary, we shall be honest about the challenges we face, and put the best foot forward.

Professor Stuart Russell, Professor of Computer Science of UC Berkeley, emphasised in his latest podcast, The promise and perils of AI' with the World Economic Forum[156], his three principles for human centric AI:

1. The machine's only objective is to maximise the realisation of human preferences;
2. The machine is initially uncertain about what those preferences are; and
3. The ultimate source of information about human preferences is human behaviour.

To build ethical AI, we must first understand the diversity and complexity of human preferences and behaviour, and acknowledge that it is a journey and not a single destination.

5.6. Authors' Biographies

Christine Chow, PhD

Christine is the global head of Stewardship at HSBC Asset Management and a board member of HSBC Asset Management UK Limited. She has 25 years' experience in investment management, research & consulting, with a focus on technology and sustainability.

Her PhD thesis on responsible investment was short-listed for a United Nations award in Sweden for industry relevance and academic excellence. She is a board member of the International Corporate Governance Network (ICGN), an organization led by investors responsible for assets under management in excess of $US 59 trillion from over 39 countries and territories. She was appointed an honorary adviser to the Financial Reporting Council (FRC) Hong Kong in April 2021, and the Convenor (Chair) of the FRC Sustainability and Climate Action Task Force (SCATF) in February 2022.

She is an Emeritus Governor of the London School of Economics (LSE), following the completion of her six-year term as a Member of Court and Investment Committee (2015-2021). She was a member in the Data Governance Task Force of the UK All Party Parliamentary Group (APPG) on Artificial Intelligence (2018 – 2021) and an Adjunct Professor in Finance at the Hong Kong University of Science and Technology (2014 - 2016), where she established the first tri-modal graduate course on social entrepreneurship and impact investing, supported by funds from family foundation and the Hong Kong government.

She was named as one of the top 30 Inspirational Women in the City of London. In 2020, she won the Finance Monthly Women in Finance Award as the Investment Management Leader of the Year (Asia).

Christine is a graduate of the London School of Economics and the University of Melbourne. She completed an executive education course on financial engineering at Stanford University.

[156] https://www.weforum.org/agenda/2022/01/artificial-intelligence-stuart-russell-radio-davos

Mark Lewis, BA, LLB Hons, LLM

Mark is a senior consultant in the UK law firm Macfarlanes LLP, focusing on technology, fintech and financial services outsourcing advice and transactions.

He qualified as a barrister (England and Wales, Lincoln's Inn, Hardwicke Scholar) in 1982 and requalified as a solicitor (England and Wales) in 1991. For over 30 years, he has specialised in technology law and transactions. For 27 years, Mark was a partner in major City and international law firms, heading their intellectual property, technology, and commercial practices. From 1997 to 2004 he co-founded and led law firms associated with PwC and EY, from which he also led their technology, outsourcing and e-commerce global legal networks. For the first part of his career, Mark served in the UK government, where he developed a fascination for information technology.

In 2019, he was appointed a Visiting Professor in Practice in the Law School of the London School of Economics & Political Science, where he lectures on cloud computing, artificial intelligence and machine learning, and cybersecurity and resilience.

Paris Will, BSc MSc

Paris is the Lead Corporate Research Advisor at The Inclusion Initiative, the London School of Economics. In this role she carries out behavioural science research for corporate and executive clients. Her primary research topic is on assessing the applications of Artificial Intelligence with the goal of reducing decision-making bias.

In 2020, she graduated from University College London with a MSc in Industrial Organisational and Business Psychology on the Dean's Honour List. While studying, she also became a certified practitioner in multiple psychometric assessments for talent management.

Paris is the Co-Founder of Business Psych Bulletin, a research platform aiming to bridge the science-practitioner gap. She has past experience working in cognitive sciences where she has published research in top-tier academic journals and presented at international conferences.

Important Information

For Professional Clients and intermediaries within countries and territories set out below; and for Institutional Investors and Financial Advisors in Canada and the US. This document should not be distributed to or relied upon by Retail clients/investors.

The value of investments and the income from them can go down as well as up and investors may not get back the amount originally invested. The capital invested in the fund can increase or decrease and is not guaranteed. The performance figures contained in this document relate to past performance, which should not be seen as an indication of future returns. Future returns will depend, inter alia, on market conditions, fund manager's skill, fund risk level and fees. Where overseas investments are held the rate of currency exchange may cause the value of such investments to go down as well as up. Investments in emerging markets are by their nature higher risk and potentially more volatile than those inherent in some established markets. Economies in Emerging Markets generally are heavily dependent upon international trade and, accordingly, have been and may continue to be affected adversely by trade barriers, exchange controls, managed adjustments in relative currency values and other protectionist measures imposed or negotiated by the countries and territories with which they trade. These economies also have been and may continue to be affected adversely by economic conditions in the countries and territories in which they trade. Mutual fund investments are subject to market risks, read all scheme related documents carefully.

The contents of this document may not be reproduced or further distributed to any person or entity, whether in whole or in part, for any purpose. All non-authorised reproduction or use of this document will be the responsibility of the user and may lead to legal proceedings. The material contained in this document is for general information purposes only and does not constitute advice or a recommendation to buy or sell investments. Some of the statements contained in this document may be considered forward looking statements which provide current expectations or forecasts of future events. Such forward looking statements are not guarantees of future performance or events and involve risks and uncertainties. Actual results may differ materially from those described in such forward-looking statements as a result of various factors. We do not undertake any obligation to update the forward-looking statements contained herein, or to update the reasons why actual results could differ from those projected in the forward-looking statements. This document has no contractual value and is not by any means intended as a solicitation, nor a recommendation for the purchase or sale of any financial instrument in any jurisdiction in which such an offer is not lawful. The views and opinions expressed herein are those of HSBC Asset Management at the time of preparation, and are subject to change at any time. These views may not necessarily indicate current portfolios' composition. Individual portfolios managed by HSBC Asset Management primarily reflect individual clients' objectives, risk preferences, time horizon, and market liquidity. Foreign and emerging markets. Investments in foreign markets involve risks such as currency rate fluctuations, potential differences in accounting and taxation policies, as well as possible political, economic, and market risks. These risks are heightened for investments in emerging markets which are also subject to greater illiquidity and volatility than developed foreign markets. This commentary is for information purposes only. It is a marketing communication and does not constitute investment advice or a recommendation to any reader of this content to buy or sell investments nor should it be regarded as investment research. It has not been prepared in accordance with legal requirements designed to promote the independence of investment research and is not subject to any prohibition on dealing ahead of its dissemination. This document is not contractually binding nor are we required to provide this to you by any legislative provision.

All data from HSBC Asset Management unless otherwise specified. Any third party information has been obtained from sources we believe to be reliable, but which we have not independently verified.

HSBC Asset Management is the brand name for the asset management business of HSBC Group, which includes the investment activities provided through our local regulated entities. HSBC Asset Management is a group of companies in many countries and territories throughout the world that are engaged in investment advisory and fund management activities, which are ultimately owned by HSBC Holdings Plc. (HSBC Group). The above communication is distributed by the following entities:

- In Argentina by HSBC Global Asset Management Argentina S.A., Sociedad Gerente de Fondos Comunes de Inversión, Agente de administración de productos de inversión colectiva de FCI N°1;
- In Australia, this document is issued by HSBC Bank Australia Limited ABN 48 006 434 162, AFSL 232595, for HSBC Global Asset Management (Hong Kong) Limited ARBN 132 834 149 and HSBC Global Asset Management (UK) Limited ARBN 633 929 718. This document is for institutional investors only, and is not available for distribution to retail clients (as defined under the Corporations Act). HSBC Global Asset Management (Hong Kong) Limited and HSBC Global Asset Management (UK) Limited are not authorised deposit-taking institutions and are exempt from the requirement to hold an Australian financial services license under the Corporations Act in respect of the financial services they provide. HSBC Global Asset Management (Hong Kong) Limited is regulated by the Securities and Futures Commission of Hong Kong under Hong Kong laws, which differ from Australian laws. HSBC Global Asset Management (UK) Limited is regulated by the Financial Conduct Authority of the United Kingdom which, for the avoidance of doubt, includes the Financial Services Authority of the United Kingdom as it was previously known before 1 April 20 13, under the laws of the United Kingdom, which differ from Australian laws. An investment in the product does not represent a deposit with, or a liability of, HSBC Bank Australia Limited. HSBC Bank Australia Limited does not stand behind HSBC Global Asset Management (Hong Kong) Limited or HSBC Global Asset Management (UK) Limited, or guarantee the performance of the product;
- In Bermuda by HSBC Global Asset Management (Bermuda) Limited, of 37 Front Street, Hamilton, Bermuda which is licensed to conduct investment business by the Bermuda Monetary Authority;
- In Canada by HSBC Global Asset Management (Canada) Limited which provides its services as a dealer in all provinces of Canada except Prince Edward Island and also provides services in Northwest Territories. HSBC Global Asset Management (Canada) Limited provides its services as an advisor in all provinces of Canada except Prince Edward Island;
- In Chile: Operations by HSBC's headquarters or other offices of this bank located abroad are not subject to Chilean inspections or regulations and are not covered by warranty of the Chilean state. Further information may be obtained about the state guarantee to deposits at your bank or on www.sbif.cl;
- In Colombia: HSBC Bank USA NA has an authorized representative by the Superintendencia Financiera de Colombia (SFC) whereby its activities conform to the General Legal Financial System. SFC has not reviewed the information provided to the investor. This document is for the exclusive use of institutional investors in Colombia and is not for public distribution;
- In Finland, Norway, Denmark and Sweden by HSBC Global Asset Management (France), a Portfolio Management Company authorised by the French regulatory authority AMF (no. GP99026) and through the Stockholm branch of HSBC Global Asset Management (France), regulated by the Swedish Financial Supervisory Authority (Finansinspektionen);
- In France, Belgium, Netherlands, Luxembourg, Portugal, Greece by HSBC Global Asset Management (France), a Portfolio Management Company authorised by the French regulatory authority AMF (no. GP99026);

- In Germany by HSBC Global Asset Management (Deutschland) GmbH which is regulated by BaFin (German clients) respective by the Austrian Financial Market Supervision FMA (Austrian clients);
- In Hong Kong by HSBC Global Asset Management (Hong Kong) Limited, which is regulated by the Securities and Futures Commission;
- In India by HSBC Asset Management (India) Pvt Ltd. which is regulated by the Securities and Exchange Board of India;
- In Israel, HSBC Bank plc (Israel Branch) is regulated by the Bank of Israel. This document is only directed in Israel to qualified investors (under the Investment advice, Investment marketing and Investment portfolio management law-1995) of the Israeli Branch of HBEU for their own use only and is not intended for distribution;
- In Italy and Spain by HSBC Global Asset Management (France), a Portfolio Management Company authorised by the French regulatory authority AMF (no. GP99026) and through the Italian and Spanish branches of HSBC Global Asset Management (France), regulated respectively by Banca d'Italia and Commissione Nazionale per le Società e la Borsa (Consob) in Italy, and the Comisión Nacional del Mercado de Valores (CNMV) in Spain;
- In Mexico by HSBC Global Asset Management (Mexico), SA de CV, Sociedad Operadora de Fondos de Inversión, Grupo Financiero HSBC which is regulated by Comisión Nacional Bancaria y de Valores;
- In the United Arab Emirates, Qatar, Bahrain & Kuwait by HSBC Bank Middle East Limited which are regulated by relevant local Central Banks for the purpose of this promotion and lead regulated by the Dubai Financial Services Authority.
- In Oman by HSBC Bank Oman S.A.O.G regulated by Central Bank of Oman and Capital Market Authority of Oman;
- In Peru: HSBC Bank USA NA has an authorized representative by the Superintendencia de Banca y Seguros in Perú whereby its activities conform to the General Legal Financial System - Law No. 26702. Funds have not been registered before the Superintendencia del Mercado de Valores (SMV) and are being placed by means of a private offer. SMV has not reviewed the information provided to the investor. This document is for the exclusive use of institutional investors in Perú and is not for public distribution;
- In Singapore by HSBC Global Asset Management (Singapore) Limited, which is regulated by the Monetary Authority of Singapore;
- In Switzerland by HSBC Global Asset Management (Switzerland) AG whose activities are regulated in Switzerland and which activities are, where applicable, duly authorised by the Swiss Financial Market Supervisory Authority. Intended exclusively towards qualified investors in the meaning of Art. 10 para 3, 3bis and 3ter of the Federal Collective Investment Schemes Act (CISA);
- In Taiwan by HSBC Global Asset Management (Taiwan) Limited which is regulated by the Financial Supervisory Commission R.O.C. (Taiwan);
- In the UK by HSBC Global Asset Management (UK) Limited, which is authorised and regulated by the Financial Conduct Authority;
- And in the US by HSBC Global Asset Management (USA) Inc. which is an investment adviser registered with the US Securities and Exchange Commission.

NOT FDIC INSURED / NO BANK GUARANTEE / MAY LOSE VALUE