AI INEQUALITIES AT WORK

Data Justice Lab

TUC Cymru

March 2025

AI INEQUALITIES AT WORK

Report by the Data Justice Lab for TUC Cymru

Lina Dencik, Jessica Brand,
Philippa Metcalfe and Cate Correia Hopkins

Data Justice Lab

TUC Cymru

March 2025

Acknowledgements

We would like to thank Gavin Pearce for designing the layout of this report and Alice Arkwright, Vic Jones, Nisreen Mansour, Mary Towers, Ceri Williams and Rhianydd Williams for their guidance and feedback on its content.

The graphics used in this report are for illustrative purposes only and are under license from Evanto Elements.

Table of Contents

Executive Summary

The report provides a review of research to date on AI and (in)equality in the workplace, focusing on the following areas: Age; Women; Disability; Ethnicity; and Minority Languages. Drawing from studies across the world, the report highlights that although AI is often advanced on grounds of efficiency and enhanced productivity, including more objective decision-making, the impacts of data-driven technologies, including AI, on work and workers so far tend to extend or introduce significant inequalities. By showcasing how such inequalities are present and become manifested within the workplace across different groups of workers, the report highlights both the intersectional nature of AI inequalities as well as the particularities of different lived experiences.

Looking at uses of AI during key stages of workplace relations and the labour process, from hiring and recruitment through to management including the direction, evaluation and the disciplining of workers, as well as the more recent adoption of generative AI in workplace settings, the report provides a comprehensive overview of the complex and multifaceted nature of AI inequalities. The report shows how significant inequalities are manifested within the workplace due to both the nature of the technology itself, particularly in terms of what data is or can be generated and collected, how that data is processed, and the nature of outputs and decision-making that result from such processes, as well as the importance of the broader context in which such technology is developed and used.

By exploring research on questions of age, women, disability, ethnicity and minority languages in relation to AI and inequality, the report makes clear that the disparate impact of AI on different workers is intricately linked to historical patterns of social and economic inequality that sees the already advantaged reap most of the benefits of AI whereas those already disadvantaged tend to be the most at risk of harm. Such harm can happen by being more exposed and subjected to the use of AI technologies within the labour process, by being more likely to experience discriminatory outcomes based on their use, or by being less equipped with the resources needed to exploit the opportunities of AI. For example, both women and young workers tend to occupy more precarious positions in the labour market where experimentation with AI technologies in the management of workers has become more widespread, such as care work and platform labour, often involving increased surveillance and work intensification. They also dominate in jobs that are more likely to be replaced by AI-driven automation. Furthermore, the reliance on such technologies has been shown to particularly harm older workers and ethnic minorities whose identities and experiences are not properly accounted for in the design and use of data-driven systems. Similarly, disabled workers and minority language workers are often found to be stigmatised or excluded when new technologies are introduced. At the same time, the report also showcases some of the ways AI has been used to support or advance equality through more inclusive technologies or highlighting existing discriminatory practices within organisations. For example, active efforts have been made to create new data-driven models catered explicitly to minority languages to further advance their use in society more broadly. Similarly, research shows that AI technologies can be used to actively include disabled workers in processes where they were previously excluded. The use of AI in recruitment and hiring can also seek to deliberately target historical practices that have resulted in biased or exclusionary outcomes, perhaps particularly with regards to women and ethnic minorities, and to allow for more inclusive forms of recruitment.

While such advancements are welcomed in responses to the use of AI in the workplace, the report also makes clear that overwhelmingly efforts from within and beyond the labour movement have been oriented towards minimising harms of AI, often ex post, through securing more transparency and better safeguarding measures, or by seeking to end or limit the use of AI technologies for certain purposes or in particular settings. In this sense, the report shows the continued need to mobilise efforts that can also address AI inequalities preventatively, both through stricter regulation, including avenues for AI to be refused, as well as through enhancing workers' voices and decision-making power within workplaces in ways that actively foregrounds the experiences of those workers most likely to be harmed or disadvantaged by its use.

Introduction

The rapid advancement of Artificial Intelligence (AI) across workplaces in the UK and beyond has been accompanied by a growing debate on its impact on both the nature of work and on workers. While much attention has been focused on issues of job losses due to automation or on the growth of new forms of work associated with the rise of platforms, more general concerns with the relationship between AI and (in)equality are becoming prevalent. These concerns are not necessarily new but build on long-standing engagements with the way technology is bound up with existing social structures and power dynamics that shape both its design and its implications. The accelerated assertion of AI as a cornerstone of global economies and the extensive investment in AI technologies across social and public life adds pertinence to the need for an extensive engagement with the way such a vision of AI might impact on people's working lives. A focus on AI inequalities draws attention to the way this impact may be significantly varied depending on who is the subject and the context within which technologies are being used.

In this report, we bring together research that showcases the intricate relationship between AI and (in)equality as it applies to the workplace. The focus is on five different categories of workers that individually and collectively highlight the disparate impacts of data-driven technologies: i) Age; ii) Women; iii) Disability; iv) Ethnicity; and v) Minority Language Speakers. These groups are significant not only because research suggests that they may stand to be particularly impacted by the speedy acceleration of AI across the UK economy, but also because they occupy important historical positions in trade union engagement that require special consideration. Of particular importance is the understanding of these different groups of workers as deeply connected to each other, while also presenting different challenges that lead to varied experiences of AI.

The turn to data-driven technologies, such as AI, has been shown to cater predominantly to majority experiences and often struggles to account for minority and marginalised populations. The development, design and implementation of such technologies are often contingent on existing structures in society that either exacerbate inequalities or lead to new forms of inequality when these technologies are used. For example, studies have shown how data collection often entails the overrepresentation or underrepresentation of data relating to certain groups that lead to skewed outputs or predictions. What and how data collected is significant as AI technologies tend to rely on pattern recognition based on machine learning that therefore will depend on the nature of such data. Importantly, this also means that data may not explicitly include demographic characteristics or group identities but may still result in disparate impacts on particular communities. That is, the vast range of data sources used for AI means that some of this data may also serve as proxy data for demographic features, including protected characteristics such as age, gender, ability, and ethnicity, that can unintentionally lead to discriminatory outcomes.

Furthermore, the design of the algorithms that inform how data is collected and processed, and what AI models might consist of, is often reflective of the knowledges and experiences that dominate in different cultures, including in the fields of computing and engineering. This means that although technologies are optimised to respond to real-world tasks, research has shown that such technologies may not be well-suited for complex lived realities that require types of knowledges that are not easily translated into computational systems. Instead, such systems tend to rely on abstraction and reductionism in order to allow for complex data to result in singular outputs. This abstraction and reductionism of context risks especially harming those who are already marginalised or those who do not necessarily conform to dominant understandings of lived experience.

In many cases, these concerns about how the nature of data and the design of models might have disparate impact are translated into a concern with so-called 'algorithmic bias'. This understanding of the relationship between AI and inequality tends to focus on the technical features of data-driven technologies and how these might therefore also be addressed through forms of 'bias mitigation' or 'de-biasing' algorithms. The emphasis on algorithmic bias has been prominent in AI policy debates, especially in Europe, where measures to address AI-driven discrimination through a number of technical fixes and safeguards, such as increased impact assessments and auditing, have been actively advanced (Niklas and Dencik 2024). However, as this report makes clear, AI inequalities are not confined to technological features but instead encompass the broader political, social and economic dimensions within which AI technology is embedded. This is important as it showcases how an engagement with AI inequalities requires us to consider not only the technology itself, but perhaps more significantly, the contexts within which it is developed, used and experienced.

The rapid implementation of AI technologies at work provides a particularly pertinent setting for considering AI inequalities that highlight how such inequalities are bound up with different histories, practices, and social relations. The growing reliance on data-driven technologies across key stages of the labour process, starting with recruitment and hiring, and continuing to management, including the direction, evaluation and disciplining of workers, has been shown to significantly impact on issues of (in)equality. With the advent of not only analytical AI but also generative AI in the workplace, some of these issues are becoming further heightened. Importantly, the introduction of new technologies in workplaces is often linked to increased managerial control and pursuits of enhanced productivity, as well as claims to more objectivity and less bias in decision-making and treatment regarding workers and the possibility for more technology-assisted inclusive practices. In this sense, the relationship between AI and inequality as it relates to work and workers is multifaceted and requires an engagement with the many complexities of socio-technical systems as they exist in the real world, including their potentially contradictory impacts.

In this report, we set out a comprehensive picture of research to date on the relationship between AI and (in)equality as it relates to the workplace, focusing on different connected groups of workers. For each section engaging with a different group, we focus on key areas of research relating to different aspects of human resources and workplace management, including i) Hiring and Recruitment; ii) Algorithmic Management; and iii) Generative AI. Each section then ends with a review of relevant responses to AI inequalities or ways in which equality might be advanced with AI. The section on minority languages follows a slightly different structure, reflecting the nature of research available about that particular topic, exploring especially how AI technology may or may not support minority language use within and beyond the workplace. Crucially, although the report treats each demographic group as a discreet group, it will be apparent throughout the report that in many cases experiences of inequality are compounded by the intersectional nature of AI's impact in the workplace. As such, while AI may impact almost all workers, it is often those with protected characteristics who are most vulnerable to the risks that come with that impact. It is therefore important to engage with the report as a whole at the same time as recognising some of the particularities that apply to the working lives of different groups of workers.

Overall, the report illustrates the continued need to engage with AI's impact on working lives, especially amongst those that occupy already precarious or marginalised positions in the labour market. By highlighting a multitude of case studies and research from around the world, the report provides concrete evidence for how AI tends to extend and entrench inequalities in the workplace despite possibilities for also empowering and supporting otherwise disadvantaged workers. The report is therefore a call to action for more direct engagement with AI inequalities, building on efforts such as the TUC's AI Bill and the more targeted efforts from within and beyond the labour movement outlined within the report. As the Welsh Parliament alongside the UK government set out action plans for an increasingly AI-driven economy, it becomes more important than ever for those who represent and seek to advance workers' interests to ensure that such plans consider their implications not just for the majority, but for *all* workers.

AI Inequalities: Age

Summary

Inequalities relating to AI and age are evident across processes of recruitment and hiring, management of workers, and the more recent arrival of generative AI in workplaces. Age-related inequalities are present for both young workers and older workers. Young workers tend to be disproportionately exposed to and impacted by AI as uses of AI for the purposes of hiring and management are overrepresented in entry-level jobs and low-wage and precarious work that tend to have a larger proportion of young workers. At the same time, older workers experience exclusion and harmful impacts with uses of AI as they are either unable to access employment or are 'managed out' through algorithmic management techniques. Although engagement with AI might be most prevalent amongst 'prime-aged' white collar workers between 25-54, these workers tend to be better positioned in terms of finding such tools complimentary to existing skills rather than requiring upskilling or potential replacement. Such divisions look set to be widening with the uptake of generative AI by workers declining with age, and older workers are also more likely to experience emerging technologies as negative or harmful. Efforts within the EU and elsewhere have therefore been made to advance more age-sensitive design for work environments and calls have been made in the UK to account better for older workers in digital inclusion policies and to provide more digital skills education for younger people.

Introduction

The differential impact AI might have on different age groups features as a prominent concern in debates on the future of work, but the relationship between AI, age and employment has been the subject of limited research in and of itself in AI and remains fragmented (Komp-Leukkunen et al. 2022). This is despite long-standing research showcasing generational differences in attitudes, uses and adaptation of new technologies and demographic shifts relating to age within sectors most likely to adopt new technologies. The IMF, for example has highlighted that 'as AI reshapes the labour market, workers will likely adapt to shifting demands, with outcomes varying by education and age' (2024: 15). Digital divides pertaining to access, skill and literacy have been widely evidenced to particularly enhance the vulnerability of older workers who are more likely to consider themselves to have low digital skills and therefore more afraid to lose their jobs and find it harder to gain reemployment (Vasilescu et al. 2020, IMF 2024). This 'grey digital divide' is likely to increase with AI as older adults become more vulnerable to the distributional impacts of digitalisation across societies (Ferdous 2023).

At the same time, young people tend to be disproportionately exposed to and in competition with AI with entry-level positions most likely to be at risk of being displaced by generative AI and sectors with high levels of automation, such as platform labour, overwhelmingly made up of young workers both in the UK (MacDonald & Giazitzoglu 2019) and globally (IMF 2024). Despite assumptions about adaptability amongst younger generations, sometimes referred to as 'digital natives', young workers also paradoxically find themselves at the fringes of the global labour market due to wider labour market structures that mean they are more likely to be in insecure, low paid work (Burgess and Connell 2020). Moreover, young people's ability to adapt to continuous mobility across different jobs, and different digital skills, is highly contingent on other demographic categories than age, such as education, nationality and economic prosperity that create unequal experiences of AI amongst workers (Zur and Zur 2011).

This also means that although studies have predicted that 'prime-aged workers', aged 25-54, with tertiary education in better-paid 'white collar' jobs, such as health care professionals and legal professionals as well as programmers and engineers, are most likely to see a big impact from AI on their job roles (Muro et al. 2019), they are significantly differently positioned in relation to what the consequences of this impact might be. In their exploration of the impact of AI on US workforces, Muro et al. (2019) argue that, unlike concerns about displacement or lack of agency associated with automation, prime-aged workers with tertiary education will be 'disproportionately involved with AI', but are more likely to be able to adapt to this exposure. This also extends to workers across the globe, as well as other demographic groups, such as migrant workers who are less likely to be in better-paid jobs with tertiary education or equivalent qualifications, according to global surveys conducted by the OECD (Lane 2024).

In this chapter, we explore the different ways in which AI inequalities have been found to relate to questions of age, starting with recruitment processes, before discussing the use of AI in management and the advent of generative AI.

AI in Recruitment and Hiring

For young people in the UK, it has been noted that the use of AI in the recruitment and hiring process is particularly pertinent as it is most widespread in the hiring of entry-level positions at large organisations where young people are overrepresented (Jung and Desikan 2024). Indeed, in a study of 'digital natives' in Nordic countries, students aged 20-23 years old saw 'AI as the future of recruitment regardless of its challenges' (Hekkala and Hekkala 2021). Similarly, in the United States, elite universities have been reported to encourage their students to adapt their performance in applications and interviews to align with expectations of AI models, such as ensuring particular keywords are used or adapting certain facial expressions (Harwell 2019). This is despite research showing that young people have concerns about how AI might be used in hiring practices and the potential for furthering bias without human input (Hekkala and Hekkala 2021).

Kim (2019), for example, has highlighted how historically discriminatory data can lead to harmful outcomes that embed bias. One aspect of this may be the use of proxy variables as stand-ins for legally proscribed criteria for hiring and the use of proxies to sort candidates implicitly on the bases of a protected characteristic (Kim 2019, Ajunwa 2020). This can occur intentionally if the designer of the algorithm knows that 'a certain trait is correlated with a protected characteristic and uses it to screen out a disfavoured group' (Kim 2019: 7). For example, research has shown that older workers are often disadvantaged when applying for jobs due to 'explicit age stereotypes' including those relating to a lack of digital skills (Zaniboni et al. 2019). Furthermore, implicit age bias, or ageism, has been shown to be more pervasive in hiring decisions than other well-documented forms of bias, such as race and gender biases, and is also more challenging to mitigate against in AI models (Harris 2023). In a study in Sweden from 2019, researchers created 6000 fictitious resumes, where they randomly assigned information about age (between 35 years and 70 years old). They found a 'strong negative age effect in all occupations', with call back rates declining substantially as age increased. They also found that as call back began to decrease fairly early in the age range tested and then flattened out, discrimination tended to be not so much about being older (over 45), but rather about not being young (under the age of 40-45). Moreover, they also found that the drop in call backs was steeper for women applicants, highlighting how age and gender intersect in discriminatory hiring practices (Carlsson and Eriksson 2019).

Older workers may also be excluded earlier on in the recruitment process, through targeted job ads. AI is increasingly a feature of where and when potential candidates might be reached for job recruitment. A lawsuit in California in 2017, for example, included plaintiffs who alleged that companies, including T-Mobile and Amazon, targeted younger Facebook users when advertising jobs, and deliberately excluded older people. This was done, for example, by paying Facebook to only show job adverts to users between 18-30 and by using behavioural data to target 'young professionals' and placing adverts in Facebook groups likely to exclude older workers, such as having 'Millennials' in the title (Kim 2019, AARP 2019).

A different example is provided by Ajunwa (2020) who discusses a case from the US of a man in his 40s who is not able to finish a job application as the automated options on the application form do not allow him to provide the year he graduated from college, with the options only beginning in 1995. Practically, this means the exclusion of all applications over a certain age. Importantly, Ajunwa argues that in these cases there is often no data trail to follow as no job application was processed and it would not necessarily be captured in any later audits of the hiring system. A similar example has been highlighted by AARP (formerly the American Associates of Retired People) who reported on a lawsuit involving PriceWaterhouseCoopers in the United States, who had used automated hiring processes to remove job applications from candidates who did not have an email address ending in .edu, excluding both older workers and those without college education (AARP 2019).

AI in Management

The advent of algorithmic management, including the use of AI, is particularly prominent in sectors that tend to be dominated by young workers, most notably in the platform economy and in warehouses across Western economies, including Europe, US, and Australia (Wood 2021). The average age for platform workers in Europe, for example, is 30 for women and 32 for men (EIGI 2021). Often, high levels of automation within these sectors are closely associated with increased forms of precarity, alongside extensive surveillance and data-driven performance assessments, that mean that young workers are disproportionately exposed to the potential harms of the use of AI in management. Age is a key marker for disparate access to stable and secure jobs, with young workers' more likely to take jobs in the gig economy and e-commerce warehouses due to a lack of alternative employment options (MacDonald & Giazitzoglu 2019). Young workers are often also more at risk of becoming trapped in precarious jobs (EIGE 2021).

Extensive research has shown how sectors with highly automated management, such as platform labour and warehouse work, both shifts and extends managerial control and worker vulnerability based on various technologically mediated forms of direction, evaluation and discipline (Kellogg et al. 2020). For example, research on platform labour has showcased how customer-focused work on digital platforms make young workers increasingly vulnerable to losing access to work based on rankings as a result of customer ratings (Rosenblat and Stark 2016; Wood 2021). Workers with lower rankings or scores are often excluded from shifts or tasks on digital platforms, in many cases without adequate transparency and lines of accountability in place despite research indicating that lower rankings can be linked to gender and/or racial prejudices (Bajwa et al 2018; Stark and Pais 2021).

Research into young workers' experience of platform work and algorithmic management demonstrates that country contexts and labour market relations likely impact the reasons a young worker may end up or remain employed in gig work. For example, Lauresen, Neilson and Dyreborg (2021) remark that British young workers often end up in gig work roles because of a lack of alternative employment. However, they found that Danish young workers often enjoy the flexibility of the work structure as they can choose when and where they work, but that the lack of transparency around task distribution, work direction, evaluation and discipline also create feelings of 'unfairness and disorientation'. Laursen et al. refers to this tension experienced by Danish workers as the 'double autonomy paradox of young workers'. They highlight how, through algorithmic management, managers become invisible, and control becomes embedded in ways that obfuscate the underlying mechanisms of the platform that dictates their working life. This means that, compared to other cohorts of workers who are still able to interact with a human manager, young workers do not have control over work processes, such as delivery routes and rankings, or fair and equal access to support and clear instruction. Laursen et al. note that the detrimental impact this might have on the health and well-being of young workers is made worse in light of being new on the labour market, often working part-time in precarious conditions and often not well-integrated into social networks in the workplace. In a report on the health effects of gig work in Canada, Bajwa et al. (2018) argue that while gig workers share some vulnerabilities, which have important negative consequences on their health, with other workers, there are platform-specific vulnerabilities for workers that require more consideration for the future health of young workers.

In warehouse work, which is also dominated by young workers, research has similarly pointed to the use of AI to create highly controlled workplaces, including the reliance on digital tracing technologies, such as handheld and wearable devices (MacDonald & Giazitzoglu 2019). These devices direct tasks and measures performance and have been found to be used to intensify worker surveillance, curtailing autonomy or worker control over the labour process (Gent 2018, Cant 2020). Companies like Amazon, for example, use handheld devices to rank employees and carry out disciplinary actions based on performance metrics. While such techniques are most prominent in work that overwhelmingly recruits young workers, research has also shown that they are rapidly migrating to other sectors, including those that have historically included a more diverse age-range of workers. Research on postal work, for example, has shown how employers such as Royal Mail have sought to replicate uses of AI and algorithmic management, across its organisation, from sorting parcels to using tracking devices on postal workers. In this case, the reliance on such technologies has been central to 'managing out' older workers who struggle to adapt to the intensification of the labour process with the use of new technologies (Brand and Dencik, forthcoming). Often, this includes changing routes and targets for deliveries based on younger bodies that are then used to remove work from older workers.

Generative AI

The advent of generative AI in the workplace is seen to follow many of the existing patterns of age-related impacts in the labour market, with ‘digital natives’ considered to be the most impacted by its uptake and also the age group most confident in adopting generative AI into their work. In a survey from the World Economic Forum in 2024 with 25,000 working adults from around the world, in the age group 18-24, 71% of men and 59% of women use generative AI tools at least once a week. In comparison, 34% of women and 42% of men in the age group 55-65 use generative AI tools at least once a week (WEF, 2024). Research has also found that older workers are much less likely to embrace the update of generative AI and face a higher risk from its widespread adoption by lacking adequate skills (Ferdous 2023).

Unlike some of the previous iterations of automation, the IMF claim that generative AI poses a higher risk to ‘high-skilled’ workers by replacing tasks beyond mechanical and routine work, producing codes and texts, and applying reason and abstract concepts (IMF 2024). As the first phase of generative AI adoption in the workplace is oriented towards back office, part-time, and entry-level jobs, those most impacted by its update will tend towards women and young workers, including graduates (IPPR 2024). To understand further impacts of generative AI, it may be necessary to move beyond what Pizzinelli and Tavares (2024) refer to as a ‘task framework’ in order to account for additional dimensions of AI, including the social, ethical and physical contexts of occupations as well as required skill levels. In what they introduce as an ‘index of AI complimentarity’, they therefore look at where AI will complement tasks and what occupations are likely to be shielded from AI-driven job displacement. This framework is premised on generative AI being more likely to augment tasks in occupations where it is complementary, increasing productivity and support, rather than replacing tasks. Pizzinelli et al. (2023), drawing on research that explored the impact of AI on labour markets in two ‘advanced economies’ (the UK and the US), and four ‘emerging economies’ (Brazil, Colombia, India, and South Africa), argue that although they do not see a ‘straightforward association between age and AI exposure’, younger workers are less likely to be in jobs with high complementarity elements, and so are ‘more susceptible to potential negative impacts stemming from widespread AI adoption’ (p. 25).

Importantly also, as demands for AI specialist jobs increase, particularly with the impact of generative AI, high entry barriers emerge to work, creating problems for new workers to enter the labour market (Weng and Lu 2025). In their analysis of over half a billion jobs across 15 countries, including the UK, US, Canada, Australia, Singapore, New Zealand and countries across Europe, PwC found that job openings that require specialist AI skills have grown 3.5 times faster than other job opening and offer a 25% wage premium compared to other job openings. Indeed, in a global CEO survey, 69% of CEOs believed that generative AI would require their workforce to develop new skills (PwC 2024). These developments can present particular barriers to young workers looking to enter the labour market, but may be lacking in opportunities to access training schemes or learning new skills in comparison to those already in work. In the UK, a new tertiary education and skills body, Medr, was set up in Wales in 2020 to support skills training and tackle inequality (gov.wales 2020) and Skills England was launched in 2024 to actively target the skills gap, increase diversity and boost AI education across schools and universities, with the aim to 'produce more AI graduates and offer job-relevant training.' (gov.uk 2025).

At the same time, research also indicates that young workers may particularly benefit from generative AI as it can 'change the way workers perform and learn' (Brynjolfsson et al. 2023). For example, in call centre work that tends to have large proportions of young workers in the UK (UNISON n.d), AI can be used for conversational assistance, increasing productivity, especially for novice and low-skilled workers, leading to a levelling effect (see also Wilmers 2024). Tacit knowledge is made accessible to other workers through AI by embodying the best practices of high-skilled workers. Reports on the reception of generative AI also indicates that young workers look to generative AI for support when managers fall short in providing it, seeing generative AI as an avenue for accessing training and gaining new skills. For example, in a survey of young workers, 56% of 18-24 year olds and a third of 25-34 year olds have used generative AI for upskilling, compared to 15% of workers aged 55-64 (Consultancy UK 2024). Moreover, a report by Accenture claims that not only are British officer workers leading the way in adopting generative AI in the workplace, but that 63% of those partaking in the study experienced positive impacts on their level of job satisfaction (Brand 2023). It is worth noting, however, that research into algorithmic management has shown the central role that managers play in how the introduction of new technologies, including AI, is experienced by workers and the extent to which it is used to level skill across the workplace or whether it is predominantly used to extend surveillance systems and managerial control (Dencik et al. forthcoming). As Wilmers (2024) describes it, we are in the flux of a 'race between skill levelling and declining costs of surveillance' that could determine whether AI 'undermines or locks in intensive monitory and surveillance systems.'

Responses

AI inequalities relating to age feature in diverse and complex ways that require consideration for how these might be addressed. Ferdous (2023), for example, argues that the diverse characteristics of older workers need to be better accounted for in policy, such as the UK's digital inclusion charter produced by the Government Digital Service in 2014. This has largely ignored older adults or have lumped them together with other age groups that make existing 'active aging' policies not appropriate for an AI digital era. Instead, they argue, digital economy regulations must be updated to include 'age sensitive' strategies that enable fair and ethical digital practices at social and institutional levels oriented towards three key points: 1) how diverse groups of older workers are able to prepare themselves for AI-driven workplaces; 2) what organisational strategies are needed to ensure older workers are able to sustain their later working lives in a digital economy; and 3) which employment policies and labour market regulations will provide suitable safeguarding for older workers vulnerable to harmful impacts of AI across the labour market. These strategies must account for changing skill requirements, digital data bias, and age-inclusive practices in order to extend the length of working lives in response to an aging population in Europe (Komp-Leukkunen et al. 2022). The European Union, for example, has led the project 'Ageing@work' which creates a 'personalised system to support ageing workers (aged 50+) into designing fit for purpose work environments and managing flexibly their evolving needs' to enable work-life balance and inclusion in the workplace (Giakoumis et al. 2019). Research in the UK is also actively being pursued to identify the skills needs for young people, before they enter the labour market, with two projects awarded by the British Academy in 2025 set to explore skills gaps across regions and demographics in the UK and to explore a shared vision for AI skills in practice (British Academy 2025).

Beyond regulation and policy efforts, workers themselves have also sought to resist harmful practices relating to AI, both formally and informally. This is particularly pertinent in sectors that are dominated by younger workers who tend to be less likely to be members of a trade union. In his research on Amazon warehouses in the UK, for example, Gent (2018) identified three forms of worker resistance: a) accidental, which could look like technological malfunctions that present a blockage to the functioning of the productive process; b) formal resistance which refers to acts which may not be official, but which are intentionally political and generally drawn from the historically established repertoire of organising tactics; and c) informal resistance which includes acts such as lying, intentional mistakes, obstruction, and wasting time, ranging from the mundane to the extreme. Indeed, several researchers have highlighted how workers in workplaces with extensive algorithmic management find ways to 'game' or 'trick' algorithms to regain agency, or building their own algorithmic literacy (potentially using generative AI) to ascertain autonomy (Celentano 2023, Chen 2018, Laursen et al. 2021).

Others have pointed to more organised forms of resistance, such as the host of actions taken by platform workers, including Deliveroo riders and Uber drivers, who have sought to organise through demonstrations and wild cat strikes to improve conditions, both without and together with unions, most notably the IWGB and the GMB (Gent 2018, Cant 2020). This has also included a number of legal actions, such as the entitlement to full workers' rights, including sick pay, breaks and guaranteed minimum wage (GMB 2021). However, outside of these focused actions on particular platforms, it is important to consider the age demographics of union members, with only 3.7% of union members aged 16-24, and only 20.9% aged 25-34 (Department for Business and Trade 2023). This means that AI inequalities relating to age present particular challenges for organised labour that may mean that important generational issues are not considered in responses, despite younger workers being disproportionately impacted by AI in work.

To adequately engage with age-related AI inequalities there is therefore a need to consider both how some groups of workers are particularly exposed to AI, providing both advantages and disadvantages, as well as the way AI may be particularly harmful to other groups of workers, depending on the intersection of age with other key demographic categories. In this sense, research suggests that responses to AI inequalities relating to age may need to be context-specific rather than adopting a blanket approach.

AI Inequalities: Women

Summary

The structural inequalities that pertain to gender and employment permeate the development and adoption of AI in the workplace. AI has been found to perpetuate gender stereotypes and advance gender norms that disproportionately disadvantage women. At the same time, women are also more likely to experience the increased surveillance and privacy infringements associated with uses of data-driven technologies in the workplace as harmful. Across recruitment and hiring and within the management of workers, women tend to experience forms of exclusion and discrimination from the use of AI tools that are often designed by and for men. Although there has been a proliferation of technologies explicitly designed for women, particularly within wellness, these are premised on the possible exposure of very intimate data that may be misused by employers. Female-dominated jobs may also be more at risk of being replaced by AI at the same time as women have been found to be less willing to embrace generative AI at work. Inequalities relating to AI and women can be a particular challenge for unions as women tend to be in employment that is less likely to have union representation and may therefore be more difficult to reach. Calls have been made to enhance the diversity within the education of STEM subjects, including computing and engineering, and to address gender disparities in the technology industry. Policy and regulatory efforts have also sought to highlight issues of gender stereotypes and bias as part of a broader 'Responsible AI' agenda.

Introduction

The impact of AI on gender disparities has, from the outset, been a major concern, raised as a central risk early on by key international bodies including UNESCO and the OECD. When the European Commission set up their High-level Expert Group on Artificial Intelligence, women were noted as a particular group at risk of disparate impact and exclusion (Gomez-Herrera and Koeszegi 2022). The European Institute for Gender Equality (EIGI) has highlighted how technologies such as AI will tend towards amplifying gender inequalities based on under-representation, bias and discrimination in algorithmic processes (EIGI 2021). What is more, gender stereotypes and gender-based violence have been found to be embedded within and reproduced in AI technologies, such as in the case of AI assistants that tend to 'exhibit female features, are depicted as helpful and pleasant, and perform secretarial tasks traditionally assigned to women.' (ibid.).

With regards to employment and labour, questions of gender have a long history in how technological disruptions manifest in the workplace, dating back to the mechanisation of textile work in the 18th century that pushed women out of participating in the labour market (Jung and Desikan 2024). It is, therefore, perhaps unsurprising that the way AI might structure labour markets and impact on gender equality has been raised as a key question across international and national bodies (ILO 2019, UNESCO 2022, WEF 2021). This includes an engagement with the way existing inequalities, including opportunities for work, position, status and treatment in the workplace, sexual harassment, and visibility and engagement with science, technology, engineering, and mathematics (STEM) fields plays out in the context of AI. Indeed, Gomez-Herrera and Koeszegi (2022) argue that existing gender inequalities, including stereotypes and gendered distribution of work and pay, is a 'systematic component' intrinsic to social structures that determine the diffusion and generation of inequalities on multiple levels. For example, they highlight how stereotypes and societal inequalities relate to early segregation in education systems, limiting the number of women pursuing STEM subjects and entering the technology industry, with only 36% of the global population enrolled in STEM education and 29% in ICT being female compared to 70% in health and welfare. In Europe, the proportion is even smaller with only 34% female in STEM and 17% in ICT education. This leads to a disparity in women's representation, renumeration and promotion, further impacting the retention rate of women in technology-related labour fields.

Such statistics are significant, as AI has been shown to be significantly shaped by the environment in which it is being designed, including the values and experiences of developers and engineers who risk neglecting the needs of diverse users and can further entrench or perpetuate stereotypes and exclusions (Gomez-Herrera and Koeszegi 2022). The lack of diversity in the tech sector has been highlighted as a key concern for questions of inequality, particularly considering the dominance of a few large technology companies that predominantly employ white, able-bodied, men, including in areas of AI-related research and development (West et al. 2019). How this translates into AI adoption in workplaces is therefore a key area of concern for the potential of inequalities to be exacerbated, including the extent to which women can adapt or reskill at the same rate as male workers (Gomez-Herrera and Koeszegi 2022). Furthermore, it can have consequences for the nature of gender classifications embedded within computational systems that tend towards reductionist binary categories that exclude those that do not fit within such categories (Costanza-Shock 2018).

To explore these questions further, we go on to outline key research on how AI inequalities relate to gender in key stages of the labour process and AI development, starting with issues of recruitment and hiring before looking at the use of AI in management, and the advent of generative AI. We conclude with some reflections on initiatives and responses seeking to address the intersection of gender inequality and AI.

AI in Recruitment and Hiring

The structural inequalities that pertain to gender and employment have the potential to create a 'vicious cycle' of digital inequality that impact women from the outset. Research has shown, for example, that AI products do not function in the same way across genders, such as voice recognition systems in healthcare products, that often underperform for women in comparison to men. In many cases, these exclusionary aspects intersect across demographic classifications, most notably race and gender (Gomez-Herrera and Koeszegi 2022, West et al. 2019).

In recruitment and hiring processes, the systematic exclusion or underrepresentation of women in both datasets and in the design of algorithms can have significant discriminatory effects. For example, Maliki and Naji (2024) have shown that existing societal gender inequalities will teach machine learning algorithms that women are paid less than men, that men are more likely to get business loans, that men are more likely to occupy higher status positions, and that men are more likely to get promoted. This in turn leads to biased predictions that inform decisions relating to who might be employed and on what terms. Indeed, research has outlined how automated hiring systems are used not just to recruit and filter candidates, but to profile and predict what terms might be acceptable to a candidate, including pay (Sanchez-Monedero 2018).

The prevalence of structural gender inequalities that come to be replicated in AI systems is particularly pertinent for women as the use of such technologies for recruitment and hiring are most prevalent in retail and low wage markets across the globe (Anjunwa 2020). In the UK, where the ratio for retail workers is 58% women to 42% male, women are thus more likely to be impacted by AI hiring tools (the Retail Appointment 2024). This means women are disproportionately exposed to the potential risks of AI in recruitment and hiring. At the same time, because historical data on previous employees is often used to optimise systems to identify 'fit' between an employer and candidate, where women are already underrepresented, they risk being further disadvantaged or excluded in the selection process (Kim 2022, Dencik and Stevens 2023). This may be perpetuated by the advent of relying more on personality profiling and behavioural data, such as cognitive skills, facial expressions or vocal tone, in algorithmic assessments of candidates, or data that relies on social media behaviour and risk taking (Dubber et al. 2020). For example, research has found that on platforms such as LinkedIn, men are more likely to use inflated titles and descriptions of their job roles, which also aligns with historical tendencies for men to report having more skills than they do when applying for jobs (Mohr 2014, Young et al. 2021).

An infamous example highlighting the use of AI hiring tools is the resume screening tool that was trialled by Amazon in 2018 and was found to systematically give higher scores to white male applicants. The algorithm had been fed on historical data relating to job performance using past resumes to highlight key terms, which told the algorithm that men were better performers in the warehouse, and male candidates were therefore seen as preferable to female ones (Maliki and Naji 2024). This was because the algorithm had been trained on a majority of male resumes, which caused the system to 'designate the male candidate's resumes as the norm' (Ajunwa 2020: 3). When gender was excluded as a variable, the tool still disadvantaged women by scoring female attributes as less desirable, e.g. studying particular courses such as 'Women's Studies' (UNESCO 2022). The algorithm was scrapped once the discriminatory outcomes had been revealed following revelations about the trial from a whistleblower, but the example has been used to illustrate the ways in which inequalities might manifest in AI systems, even without such intentions.

With the advancement of biometrics, such as the controversial use of facial recognition systems to assess personality traits based on facial expressions in interviews for example, these issues have taken on further relevance. Prominent research has demonstrated that these systems tend to underperform on particular groups of people, most notably Black, Asian and Minority Ethnic women (Buolamwini 2018) and struggle to account for cultural differences across global populations (Sanchez-Monedero et al. 2020). Indeed, the EU's AI Act prohibits the use of AI-based emotion recognition systems, often based on facial expressions or other biometric data, for the purposes of recruitment and hiring (EU AI Act 2023). Even as these technologies draw on more data to become more efficient, concerns remain about the scientific validation behind the assessment of personality profiling based on behavioural and biometric data that underpin the advancement of such tools in hiring, including the ways in which such an assessment will tend to advantage some and disadvantage others (Sanchez-Monedero & Dencik 2021, Dencik & Stevens 2023).

Moreover, even before the selection process, research has shown that inequalities permeate exposure and access to different parts of the labour market. Multiple studies have highlighted the way job advertisements align with stereotypes and discriminatory data. In a joint report by UNESCO, OECD and the Inter-American Development Bank (IDB), for example, they point to multiple aspects that impact the positions that women see advertised when searching for jobs online, including the use of gendered language for advertisements, the impact of which remains under-researched. Others have highlighted ways in which certain gender groups can be excluded from viewing jobs online and how major platforms such as Facebook deliver job adverts in a manner that aligns with gender and racial stereotypes, such as jobs in the STEM industry being shown to white men, jobs for cashier roles at the supermarket to women, and jobs as taxi drivers to Black men (West et al. 2019). An early study demonstrated this by showcasing how setting a user profile to 'female' on job advertisement pages resulted in fewer higher-paying jobs being shown in comparison to 'male' users (Datta et al. 2015).

Despite these concerns, automated hiring systems, including those that rely on AI, have seen a rapid uptake, often to particularly target issues of bias that are known to be prevalent in hiring processes. In several cases they are introduced into organisations to target discriminatory practices and find ways of making assessments about potential candidates that bypass historical practices of recruitment and hiring deemed problematic. For example, in the report by UNESCO, OECD and IDB (2022) they see potential in AI's ability to be trained in order to flag discriminatory patterns in language used in job advertisements in order to adjust the language. It can also be used to diversify the pool of candidates in interviews in traditionally male dominated industries and some argue that it is a helpful tool to mitigate 'unconscious bias' by reducing the impact of human decision-makers and removing demographic factors such as gender classifications in assessments that do not correlate with merit or success (Houser 2019). In these instances, it becomes important to scrutinise the basis of such claims and investigate how 'debiasing' around gender (and other protected characteristics) is carried out and with what results (Sanchez-Monedero et al. 2020).

AI in Management

In the same way as in hiring, the introduction of AI into management practices can be targeted at minimising discriminatory outcomes that organisations may have been challenged with, such as seeking ways to address gender pay gaps or inequalities in promotion processes. For example, stereotypes around women being seen as 'aggressive' as opposed to 'leadership material' when displaying similar qualities to men may be less prevalent in AI-driven performance assessments and employee reviews (Houser 2019). However, at the same time, these historically prevalent disparities between the profiling of men and women continue to inform how AI is advanced and adopted, highlighted in the predominantly female voices of AI assistants (Kohlrausch and Weber 2019). Moreover, concerns have been expressed about the extent to which existing stereotypes and expectations might feed into AI systems in such a way that further disadvantages women. For example, international bodies have warned that metrics of 'success' or 'productivity' risk being impacted by how gendered societal norms shape behaviour so that algorithmically processed assessments of customer calls in a call centre, for instance, will favour men if the standard of success is based on traits such as assertiveness or confidence that men have 'traditionally been taught to adopt in society' (UNESCO et al. 2022: 60).

We also see such biases in the growing adaptation of wearable technology in the workplace as a means of directing and monitoring workers, including for wellness. Brown (2021), for example, explores how the use of wellness monitoring tech and collection of data through wearables is used as a means of monitoring workers and improving efficiency in the workplace. She discusses the use of StrongArm, a wearable monitoring device the size of a smart phone that is used to track workers' movements that has been used by Amazon, Wallmart, Heineken and Toyota in the US alongside a plethora of other wearable devices that employers are adopting across the US, from sensor-enhanced jackets to smart glasses, virtual reality headsets, smart clothing and even mood sensing sweaters which are used to collect data on workers' movements and actions in the workplace. The disparate impact these technologies are having on workers is, according to Brown, directly linked to the lack of diverse input in the design and testing of these wearables, itself a result of embedded discrimination and exclusion of women in design and technology spaces. This has led to the roll out of wearables that do not fit women's bodies properly, or smart glasses that cannot accurately track women's gaze as well as gender-linked differences in their collection and interpretation of data.

Such discrimination may also extend to healthcare related data collected through wearable devices. Brown notes that such data is likely to ignore the fact that women have higher rates of work-related stress, anxiety and depression, or that such health factors are more likely to have more of an impact of women's long term health, including heart disease, chronic lung disease or depression, leading to an underrepresentation of women's health issues in workplace-based wellness assessments. This follows a long-standing pattern of a lack of inclusion of women, and in particular Black, Asian and Minority Ethnic women, in healthcare data collection and testing, pre-dating digital health and wellness devices in the workplace (Whelan 2021).

In part in response to this exclusion of women in health-related data, there has been a rapid proliferation of so-called 'FemTech' often designed and built by female led start-ups, focused on health issues that impact women and people assigned female at birth, including fertility issues, menopause, pregnancy, and periods. However, Brown (2021) highlights their paradoxical impact within a workplace setting, sometimes contributing to sexual harassment and discrimination against women based on technology designed to reverse gendered discrimination and bias in the workplace. For example, in the context of the US, she highlights period tracking apps such as MyFLO which advises women what they should be doing at work according to their period cycle, including when a woman might be better suited to doing research and brainstorming activities as opposed to giving a presentation if she is on her period. This is marked on the apps calendar, which might be visible to the employer and could lead to the employer refusing to let a female employee give a presentation that may be important to the advancement of their career. Another example given by Brown is the Kindara and Ava wearable bracelet used for fertility tracking. She notes that employers will likely assume anyone wearing one is looking to conceive and react to embedded biases against pregnant women in the workplace, leading to potentially 'adverse employment decisions'.

As such, the proliferation of new technologies that can collect health-related data might fill an important gap in representation for women, but in reversing this gap, women are also put at greater risk of infringements upon privacy and exposure. As Brown asks, why should women have to share their data on their sex drive or periods, or disclose when she is having fertility issues? Beyond this, with continued systemic sexism in the workplace, Brown highlights how employers may consciously or unconsciously have stereotypes in their heads that lead them to believe that women who are on their periods are unable to perform certain tasks, that women who are trying to get pregnant are not committed to their work, or that women who are pregnant may not be the best candidate for business trips, placing women at greater risk of discrimination due to disclosure of private and sensitive healthcare data (Brown 2021).

In this context, it is important to note that research has showcased the extent to which the adoption of AI in workplaces is significantly shaped by existing workplace relations and styles of management that therefore also indicates the importance of the continued dominance of men in managerial positions. For example, it is widely recognised in research that technologies such as AI have overwhelmingly intensified workplace surveillance and the monitoring of worker activity while also undermining possibilities for organising and the bargaining power of workers (Dencik et al. 2024, Williams 2024). This might be particularly disadvantageous for women who are known to experience such surveillance and monitoring as more harmful and intimate than men and may therefore also be less likely to make themselves visible. Ajunwa (2020), for example, makes the case that pervasive surveillance in the workplace and the use of workers' data for performance indicators is intimately connected to broader questions around bodily autonomy and personhood that have been central to women rights' campaigns. Furthermore, UNESCO et al. (2022), who conducted a global study of the impact of AI on the working lives of women, found that female workers are more likely to be concerned about privacy, including working from home and the risk of exposing family environments and children, for which women are disproportionately more responsible.

These concerns also extend to issues of sexual harassment regarding an uptake in remote working enabled by digital technologies but not necessarily accompanied by channels to report or address exposure to unprofessional behaviour. TUC research from 2016, for example, has showcased the unequal levels of harassment in UK workplaces with 64% of 18-24 year old women having experiences of harassment in the workplace (TUC 2016). In online environments, women are exposed to further avenues for harassment with less recourse to support or means of challenging problematic behaviour (Eisenstadt 2022). An increasingly prominent form that digitally-enabled sexual harassment in the workplace can take is through image-based sexual abuse, which research has shown disproportionately impacts women in relation to their job prospects, security and well-being. Although significantly under-researched, Rood and Schriner (2020) explore how the digital distribution of image-based sexual abuse in the workplace can seriously 'smear' a person's reputation at work, placing their job at risk, and potentially impact future career options. They argue that issues such as victim blaming, rape culture and the normalisation of general violence towards women, make it difficult for this kind of harassment to be adequately addressed. Furthermore, with the rise of 'deep-fakes' enabled by AI technologies that refer to 'machine learning-based software tools that produce realistic synthetic media content', image-based sexual harassment in the workplace has increased, including 'non-consensual intimate deepfakes' (Lafier and Rehman 2023, Kira 2024). The reesarch initiative Deeptrace (2019) have found that deep fakes almost exclusively target and harm females and the UK's Online Safety Act 2023 (OSA) outlines such type of harassment as a 'priority offence'. However, Kira (2024) notes that the OSA does little to empower survivors of such offences or prevent their dissemination in the first place as there is a lack of consistency in the mechanisms used to remove harmful content and a lack of clear routes to seek redress after harm has been inflicted.

At the same time, AI has also been used as a means to detect and report sexual harassment at work, such as the rise of the #MeTooBots in the workplace across the US and Europe, including the UK, which is part of efforts to develop AI technologies to monitor, detect, and report forms of sexual harassment at work where this has otherwise gone unreported (Eisenstadt 2022). These efforts include the introduction of AI tools that analyse and track digital communication and automatically flag and report anything problematic to Human Resources, absolving any need for the target of harassment to take action. However, as Eisenstadt points out, issues with such tools occur because they tend to reflect societal bias regarding language so that language more commonly used by someone from a particular cultural or class background is more likely to be flagged as problematic. Furthermore, they are riddled with technical 'failure' due to the challenges of AI to grasp nuances, such as contextual clues, social and cultural signifiers, the nature of communication, and the relationship between the people communicating. These failures, Eisenstadt notes, are particularly poignant in relation to sexual harassment as sexual harassment is often very hard to detect and depends upon the context, relationship, and positionality of the people involved (2022: 376).

Generative AI

As noted above, the prevalence of existing stereotypes and gendered norms have been a significant part of debates on AI, most notably perhaps in discussions on AI assistants and robots. These concerns continue to be pertinent with the advancement of generative AI as research has shown how applications such as ChatGPT follow normative cultural signifiers, such as correcting non-gendered pronouns, ranking intelligence based on gender bias, and perpetuating gendered notions of jobs, skills and education (Singh and Ramakrishan 2023). Moreover, reviews of how generative AI might impact on job displacement and job losses have highlighted the extent to which this might align with gender divisions in the labour market. While predictions about such shifts in the labour market remain speculative, research in Germany, for example, shows how men face a higher risk of displacement due to automation across all levels of qualification and work, whereas sectors such as social care and cultural jobs are less at risk of automation. On the other hand, these sectors may be at risk of devaluation due to their separation from digital skills (Kohlrausch and Weber 2020). However, Wilmers (2024) points out that with generative AI, higher paid occupations may be exposed to more risk, particularly with developments relating to Large-Language Models (LLMs) that are increasingly used across non-routine cognitive work such as writing emails, coding, and designing. Indeed, significant concerns have been expressed from the creative industries about the impact of generative AI on job losses, particularly in areas where women are dominant in the workforce (Leslie et al. 2025). Research has also highlighted that across the global labour market, roles such as clerical work, human resources, retail jobs, call centres, and banking – roles that are sometimes referred to as 'pink collar' jobs as a reference to historically female occupied positions – are at a higher risk of automation (Lawrence, 2018; Schmidpeter and Winter-Ebmer, 2018; Brussevich et al., 2019). This is pertinent where, as Dalingwater (2018) discusses, pink collar jobs are not something consigned to history, with an increasing feminisation of certain sectors continuing across France, Canada, the UK, and the US. It is also noteworthy that research into the impacts of generative AI on jobs have overwhelmingly been carried out by female researchers, breaking the pattern of STEM-related research more broadly, which may be indicative of its gendered distribution and impact in society at large.

Generative AI is increasingly being adopted to directly replace such jobs. For example, in the UK, the National Eating Disorders Association (NEDA) recently replaced their entire human helpline workforce with chatbots (Jung and Desikan 2024). Women may also be more at risk of being replaced as it has been shown that not only are many 'pink collar' jobs at risk from AI, but women have been found to be less willing to embrace generative AI at work, with 71% of men aged 18-24 compared to 55% of women globally saying they use generative AI at work each week (WEF 2024). However, how the impact of generative AI will impact workplaces in practice remains unclear. In their report based on an assessment of 22,000 work tasks across the UK economy, Jung and Desikan (2024) outline what they see as four phases of generative AI in the workplace. Phase 0 refers to experimentation and platform investment with only small scale use cases. In the first phase of generative AI deployment, back office jobs, more likely occupied by women, face increased risks. This, in turn, is likely to impact gender inequalities in the workplace more broadly that may also have consequences for subsequent phases as generative AI adoption migrates to 'white collar tasks' in phase 2 before processes start getting built around AI in phase 3 with tasks being transformed in ways we are not currently able to envisage. Importantly, however, they point out that generative AI is likely to be rolled out much faster than previous technological advancements and transformations which may make impacts greater without regulation and other societal measures to address risks properly in place. In this sense, Jung and Desikan argue that the future of generative AI and its impact on the workplace depends largely on the ability of policy makers to act swiftly to ensure it leads to higher quality, valuable, and more productive outputs by workers rather than replacing them entirely.

Responses

As AI adoption advances rapidly across labour markets, gender inequalities are key to understanding how and with what implications this advancement is happening. Existing inequalities and new ones permeate the nature of adoption and how workers experience the realities of AI. With the rapid escalation of generative AI, these issues are possibly more pertinent than ever, especially as technological advancements have not been matched with appropriate channels to report, challenge and redress disparate impacts on women in particular. There have therefore been calls to ensure that the industrial strategy being put forward for AI in the UK is sufficiently 'job-centric' in a way that 1) protects existing jobs and ensures gains for workers; 2) boosts the creation of new tasks and support job transitions; 3) addresses the fallout from lower labour demand (Jung and Desikan 2024).

Alongside this, the area of 'Responsible AI' has grown rapidly as part of delivering the government's missions in the UK and also aligns with priorities set out in the EU and elsewhere. A key aspect of this is finding ways to concretely address gender bias and inequality in the workplace based on greater transparency and accountability, and more inclusive approaches to the advancement of AI (Gross 2023). Such approaches include a focus on ensuring gender equality when training and developing algorithmic systems as a means of ensuring inclusive, responsible, and fair AI systems (Arisoy-Gedik and Ceyhan 2024). West et al. (2019) argue that this must also be accompanied by actions that work towards ending wider pay inequalities, addressing harassment, changing hiring practices and promotion pathways as a means of maximising diversity, as well as making hiring practices more transparent. For AI specifically, they stress the importance of transparency, rigorous testing across the lifecycle of AI systems, the inclusion of wider social analysis when researching bias in AI, and risk assessments to analyse whether some AI systems should even be designed at all.

However, it has been recognised that advancing a job-centric industrial AI strategy also requires an enhanced role of unions and strengthened bargaining power (Kurz et al. 2019) as does the importance of ensuring worker voice within the workplace to avoid changes being left to employers or providers of AI systems (TUC 2020). Indeed Wilmers (2024) suggests that (generative) AI can drive a 'new wave' of collective action due to the need to bargain for collective goods. However, existing gender inequalities in organised labour may also be replicated in relation to new technologies. For example, Kohlrausch and Weber (2020) highlight that in the European platform economy, existing disadvantages for female employees remain, in part, because there is less institutional and union representation and bargaining power in female dominated work across European countries. As such, they ask whether female workers will be able to gain much needed and decisive bargaining power, and whether unions and workers will even be able to 'create new tools of collectivisation under the conditions of digital labour' (p.26).

These arguments point towards a recognition that to tackle gender inequalities in AI, there is a need to engage with ways to address structural and societal inequalities more broadly. This also means going beyond diversity policies that have dominated a lot of discussion on AI, such as greater inclusion of women in STEM and AI jobs as a way to address algorithmic bias. This may go some way in ensuring gender inequalities are foregrounded more in the design and adoption of AI, but it does not address wider issues around workplace cultures, power asymmetries, harassment, hiring practices, unfair compensation, and tokenisation (West et al. 2019). Doing so requires a wider approach that can also account for intersectionality in the 'inherent and inherited bias' of technological systems today (Devlin 2023).

AI Inequalities: Disabilities[1]

Summary

AI inequalities relating to disabled workers combine issues with both overexposure and invisibility within datasets that can disadvantage their ability to get or keep a job. Research has shown that in hiring, disabled candidates are systematically excluded from AI-driven processes that rely on streamlined or limited forms of experience and are often disadvantaged by the ableist attitudes within the design of assessment or profiling tools. Such inequalities are particularly stark in speech and facial recognition tools that are unable to account for different characteristics that may present more in disabled people and are also reflected in the way ableism is encoding into classifications of 'fit' or 'success' in an organisation. Certain tools and platforms adopted in workplaces have also been found to be inaccessible to some disabled workers and the use of algorithmic management techniques can conceal some disabilities while shifting accountability away from employers and management for the harmful impacts that may result. The advent of generative AI has been found to further risks of exclusion and discrimination against disabled workers as ableist bias has been found to be prevalent in tools such as ChatGPT, although it may be less likely to feature in current auditing processes. Calls have therefore been made to improve AI auditing to better account for experiences amongst disabled people and to ensure greater responsibility on the part of employers to ensure tools do not harm or disadvantage disabled workers.

Introduction

Inequalities pertaining to AI are perhaps most starkly illustrated with regards to disabled people that bring together both issues of underrepresentation and overrepresentation in data simultaneously. The use of AI tools in work is entangled with social structures that exacerbates and perpetuates longstanding inequalities and injustices experienced by disabled people, as well as compounding such impact for those with multiple minoritised identities, such as disability, gender and ethnicity. As Goggin et al (2019: p.388) note, 'discrimination in favour of able-bodied people, or ableism, stems from the way society structures itself to favour certain types of bodily and personal characteristics over others', and this favouring of some characteristics over others is potentially what AI is entrenching in work. To some extent, structural impact stems from the encoding of ableist social attitudes and norms into the design of AI tools like assessment and hiring systems, and particularly with the datasets that are used to train AI models.

[1] The authors of the report have used language consistent with the social model of disability. Language used in studies cited in this report and therefore in quotes may differ from this.

This often presents for disabled people as simultaneously excessive visibility or invisibility within datasets that can disadvantage their ability to both get or keep a job.

Furthermore, the expansion of AI into hiring and management has created significant barriers for disabled people to enter and move within the labour market. In many sectors, AI hiring systems are simplifying the multiple kinds of experiences of disabled people while also discounting the structures, institutions, and environment a disabled person exists within. This reinforces both an discriminatory idea of an 'ideal type' of job candidate and also the medical as opposed to social model of disability. Experiences of algorithmic management also show that many disabled workers are having to grapple with inaccessible platforms and AI tools, particularly in gig work and crowd work. In addition, research so far suggests that generative AI is biased against disabled workers and especially autistic workers.

Such issues are particularly troubling as measures put in place to try and audit AI systems to ensure greater responsibility and fairness struggle to account for the diverse experiences of disabled people and are often not counted in standard auditing methods (Nugent and Scott-Parker 2022). There is therefore a real need to engage not just with AI's impact on disabled people, but also with the mechanisms in place to act as safeguards for any negative impacts. As we go on to outline below, AI inequalities relating to disability are prevalent throughout uses of AI in the workplace, from recruitment and hiring to management and the rapid advancement of generative AI at work.

AI in Recruitment and Hiring

The use of AI in the workplace is directly and disparately impacting disabled people by creating new barriers to them getting and maintaining a job. There has been substantial impact on disabled people in relation to the use of algorithms in recruitment and hiring processes. This includes a variety of tools and methods that pertain to a wide range of recruitment functions including candidate sourcing and engagement, candidate tracking, CV and resume screening, pre-employment assessments, and AI interviewing (Nugent et al 2020). Typically, AI hiring tools seek to identify candidates most likely to 'fit' or succeed in the workplace by comparing them to past employees or testing for personality traits associated with strong performance (Sánchez-Monedero et al., 2020, Dencik & Stevens 2023). Furthermore, vendors often claim that their algorithms are not only more efficient but also less biased than human recruiters, thereby 'allowing employers to improve hiring diversity' (Tilmes 2022: 1). However, such algorithms have been identified as perpetuating disability discrimination in what can be harmful ways. One reason for this is that they often attempt to measure and categorise aspects of the human body and thereby reify an 'ideal' type of candidate. For example, a particular AI product that has come under scrutiny is the Hirevue system which claims their 'scientifically validated' algorithms can select a successful employee by examining facial movements and voice from applicants' self-filmed, smartphone videos. However, it has been noted that this tool has the potential to discriminate against many disabled people whose impairments significantly affect facial expression and voice, such as deafness, blindness, speech disorders, and surviving a stroke (Fruchterman and Mellea 2018). This is particularly problematic as HireVue have said that 'facial actions can make up 29% of a person's employability score' (Engler, 2019). It should be noted that this particular component of HireVue's product was subsequently removed following criticism, but controversial facial recognition technology has been used by other providers of AI hiring tools for the purposes of profiling candidates during interviews.

Detailing some of the concerns about algorithmic hiring processes, Tilmes (2022) outlines several harms arising from how algorithms are used in hiring, creating additional, ableist barriers. Firstly, he notes that emotion analysis algorithms often misinterpret facial expressions of disabled people such as those with cleft lip or palate, achondroplasia, Down syndrome, and Parkinson's disease 'since they are omitted from training datasets'. Secondly, he observes that the gait and gesture processing software used in automated interviews are more error-prone when assessing the movements of people with tremors or amputations, while speech recognition AI tools also struggle to process the words of deaf people and those with speech impairments. Additionally, Tilmes notes that algorithms which assess sentiments in natural language 'often rate the mere mention of disability as negative and toxic'. Finally, Tilmes highlights that while studies show that disabled people perform similarly on many aspects of employability, automated interviews are often weighted towards easily measured factors 'such as a measured tone and eye contact'. As such, Tilmes argues that 'assessing applicants primarily in terms of abstract, quantitative factors does not preserve objectivity so much as obscure the underlying assumptions of AI designers' (2022: 7). Alongside this, Nugent et al (2020: 11) highlight that gamified assessments raise additional concerns related to dexterity, vision impairment, and response time. This is because games often involve tasks that are 'assessed based on speed of reaction to prompts and precision of responses, which may affect people with motor limitations, who need extra time or assistance to complete dexterity tasks.' Moreover, they note that people with visual impairments 'may require magnification and colour adjustment and additional time' while 'people with cognitive diversity may require language adjustment and additional time to read prompts'.

The growing use of AI and software in personality tests within hiring processes has also proved harmful for disabled people. Again, this is reifying an 'ideal' type of candidate based on ableist ideas of 'success'. It is noted that, as these kinds of tools operationalise certain personality characteristics as indicators of job success for specific kinds of roles (despite no proven correlation with performance), they tend to disproportionately screen out disabled people and especially autistic people and those with mental health conditions (Fruchterman and Mellea 2018). Other indicators used to determine probability of successful performance, such as gaps in employment, also serve as proxies that discriminate against disabled people (ibid). As highlighted by a disability rights campaigner in the US, 'Personality tests most directly get at disability in that they are designed in a way that asks questions that are trying to ferret out disability in many cases, particularly for those who have a psychiatric disability or autism' (Claypool et al. 2021: 37). What is more, it has been observed that, used in this way, AI technology in many cases has enabled the removal of the direct accountability of employers, putting distance between human decision makers and the outcomes of hiring processes. As Claypool et al note (2021), this means the way hiring algorithms are designed can diminish disabled people's chances at even getting an interview, let alone successfully obtaining a job, regardless of their qualifications.

Moreover, outcomes are heavily influenced by historical hiring decisions and so, since disabled people are twice as likely to be unemployed, they are less likely to be represented in data on past successful employees (Nugent et al 2020). This makes the prediction of the 'fit' of disabled applicants a problematic feature of AI hiring tools. Also, although there are many different forms of impairments and despite the large number of disabled people, the population is made up of many statistically small sets of people whose disabilities, impairments and conditions manifest in different ways (Givens 2020). This means the diversity of embodied experiences is not captured or represented by data and algorithms. For example, when a hiring algorithm studies candidates' facial movements during a video interview, or their performance in an online game, a blind person may experience different barriers than a person with mobility or cognitive impairment (Givens 2020). Further, as the unifying purpose of different hiring systems is that they are designed to distil the vast array of information about applicants down to a few select predictable features for the purpose of making quantifiable and easily comparable decisions, such systems are inherently limited at understanding the diversity of human experiences (Dencik & Stevens 2023). As Nugent et al (2020) point out, 'when systems need to cope with the reality of human diversity, whether it pertains to disability, ethnicity, gender, and other features, they often interpret complexity as an abnormality, or outlier'. As a result, 'predictability may come at the expense of the life chances of people with disabilities who are already faced with systematic disadvantages in securing employment' (Nugent et al 2020: 6-7).

This point is reiterated by Goggin et al (2019: 504), who suggest that the broad spectrum of disabilities 'makes it difficult, if not impossible, for programmers to account for differences in physical and mental characteristics when designing algorithms.' They further propose that if programmers cannot account for such differences ex ante, there is an inherent problem with algorithms trying to interpret the actions, behaviour patterns and gestures of disabled individuals, as there are not enough of them to be properly represented in training datasets. Moreover, there is an intersectional aspect to representation in datasets that is bound up with the issue of who 'counts' as disabled. As Tilmes (2022) notes, legal recognition or diagnoses is mediated by other biases, such as many physicians' undue scepticism about women's experiences of chronic pain, and the false belief that Black people have an unusually high pain tolerance. Tilmes therefore suggests that, because intersecting axes of oppression impact on an individual's ability to be recognised as disabled 'the few disabled people officially recognized as such in training datasets seem likely to be disproportionately wealthy, white, and male' (p.7).

Yet, at the same time, there is also disproportionate overrepresentation of particular characteristics associated with disabilities. Because vulnerable populations are often subjected to disproportionate scrutiny, their missteps are more likely to get noticed, recorded, and used to train algorithms. Tilmes (2022) gives the example of language processing AI in the US, which links the mention of disability to addiction, homelessness, and violence, 'since data about disabled people is gathered disproportionately in those settings'. Similarly, as Engler (2019) suggests, characteristics such as typical enunciation and speaking at a specific pace are qualities that might correlate with effective salespersons. Consequently, disabled people will not benefit if their qualities manifest physically in a way the algorithm has not seen in the training data: 'if their facial attributes or mannerisms are different than the norm, they get no credit, even if their traits would be as beneficial to the job'. This means that disabled people are more likely to be negatively overrepresented in training data, and, as the model treats such traits as undesired traits to receive less weighting, disabled people —like other marginalized groups—risk being excluded as a matter of course (Givens, 2020).

Finally, the structural impact of uses of AI in hiring (and management) is heightened by the fundamentally problematic way in which disability is rendered quantifiable by algorithms and abstracted away from their social context. Some argue that this means the use of AI in the workplace has had the effect of valorising the medical rather than social model of disability. In brief, the former pathologises disability, treating it as an impairment, while the latter views disability as a social construct arising from 'prejudiced attitudes and inaccessible environments that mark different embodiments as pathological' (Tilmes, 2022: 2). Tilmes (2022; 2020) has written extensively about this and, as he emphasizes, 'making sense of algorithmic discrimination against disabled people hinges on how one defines disability'. There is arguably a deep-rooted injustice with how we define and measure ambiguous criteria like 'fitness' and 'employability' because it reflects and reproduces structural ableism (Engler, 2019; Tilmes, 2022). In attempting to quantify and measure disability, disability status is often modelled as a one-dimensional variable, despite the fact that definitions of disability are fiercely contested, and that impairments are embodied in shifting ways that cannot be captured by one variable (Tilmes 2022). In this vein, Whitaker et al (2019) suggest that this also illuminates the limitations of relying on processes of 'debiasing' AI, which they suggest rest on 'limited and essentialist reading of fluid and socially constructed categories of identity' (p.11). Thus if we understand disability as a product of disabling environments, 'it is identity that can only be understood in relation to a given social and material context' (ibid). Their report cites Meryl Alper's observation that 'efforts to better include individuals with disabilities within society through primarily technological interventions rarely take into account all the other ways in which culture, law, policy, and even technology itself can also marginalize and exclude.' Ultimately, disability is made legible to AI in terms of medical diagnoses because aspects such as social context and the built environment are harder to measure (Tilmes, 2022). For example, if disabled people score lower on a given pre-employment test, 'to operationalize disability without accounting for the ableist structures at the root of that performance gap cannot help but attribute that gap to bodily failures'. Tilmes calls this 'algorithmic ableism'.

AI in Management

A key issue in inequalities pertaining to uses of AI in management is the way algorithmic management techniques create adverse experiences for disabled people, particularly in crowdwork and gigwork. Sannon and Cosley (2022) point out that disabled workers in the gig economy face several challenges around task and platform accessibility, as well as around performance monitoring and evaluation, which stem from platforms being structured in ways that penalise disabled workers. Similarly, Claypool et al (2021) highlight that how algorithms are used to manage employees poses new challenges for disabled individuals to successfully maintain a job and can shift accountability away from employers to provide workplace accommodations. Indeed, they suggest that the use of algorithms in gig work to remotely oversee and manage human workers is problematic as it can be used to hide discrimination, surveil individuals, and distance companies from the effects of their decisions. One example is Amazon's Flex program in which drivers pick up and deliver packages using routes indicated on an app and receive incentives and penalties from the app to guide their behaviour. As these kinds of AI tools are purported to foster efficiency - and often speed

- they are unlikely to accommodate for the lived experiences of disabled workers. Indeed, Amazon's algorithmic management system has been reported to fire the slowest people, regardless of the individual's impairment or access needs (ibid: p.38). Similarly, Tilmes (2022) highlights that disabled people who work slower than the average speed or have difficulty with CAPTCHAs (systems used to distinguish a human user from a robot) may be labelled as bots and are therefore barred from digital crowd work platforms such as Amazon's Mechanical Turk (MTurk). Furthermore, a study of 120 common types of MTurk tasks found that few comply with Web Content Accessibility Guidelines (Sannon and Cosley, 2022). For Sannon and Cosley, accessibility is an issue in MTurk because third party requesters design the tasks, often without accessibility in mind, and this means that 'tasks also often don't support the assistive technologies that expand disabled workers' abilities' (p. 13).

At the same time, there is also some evidence that disabled people could potentially benefit from crowdwork. Zyskowski et al (2015: 1683) suggest that the history of the struggle for equality in employment opportunities for disabled people is important in understanding how crowdsourcing can potentially offer a form of employment for disabled workers. They argue this is because the features of crowdwork may offer a unique proposition for disabled workers, such as the ability to work from home, avoid the frustrations of navigating inaccessible transportation, vary the pace of individual or multiple tasks, set a flexible work schedule, determine whether or not to reveal one's impairment or condition, and use their personal adaptive technologies. However, in their study with 24 disabled workers across four categories of gig work (delivery services; ridesharing; crowdwork; online freelancing), Sannon and Cosley (2022) found that these workers face inaccessible tasks, a lack of control and agency in the face of ability-unaware algorithms, and mismatches between customer expectations and worker abilities that lead to unfair evaluations. More broadly, they note that these kinds of challenges 'require a great deal of "invisible labor" that harms workers' health and income', often meaning that disabled workers 'may earn less than workers without disabilities' (p.4). They go on to highlight another study as an example that showed that while autistic people were able to complete most crowd work tasks, they took longer to complete than most workers, and 'longer completion times likely contribute to why crowdworkers who identify as having a disability earn $2.80/hour, versus $3.14/hour earned by workers without disabilities' (pp.4-5). However, it is worth noting that some gig companies are also actively working to hire disabled individuals. Lyft, for example, has partnered with the National Association of the Deaf to hire more deaf drivers, and Uber built a partnership with Communication Service for the Deaf to improve the experience for deaf drivers (Claypool et al, 2021). It is important to note, though, that concerns have been raised about research showing poor employment experiences of disabled people in insecure work including low pay and negative impacts on their health (Navani, Florisson and

Wilkes 2023). In addition, UK research shows that employers who encourage applications from disabled people do not necessarily provide accessible employment policies or practices to support and retain disabled people when in work (Bacon and Hoque, 2024).

Furthermore, there are also ways in which AI could enhance the diversity of workers within organisations. For example, automated team assembly systems use AI to create teams based on particular team-formation criteria and by suggesting potential teammates (Zhuang and Goggin 2024). An example is mydreamteam, a web-based tool used to facilitate team assembly developed by researchers from Northwestern University 'which uses preference matching and network heuristics to provide recommendations for team formation'. Although Zhuang and Goggin note that these tools have implications on building and designing for more diverse teams at work and consequently for disabled people's inclusion at workplaces, they also point out that research conducted by Northwestern suggests that the use of diversity parameters in these tools may pose problems. They suggest that 'making visible levels of diversity in teammate recommendation systems may in fact undermine diversity as individuals have a higher propensity to avoid having more diverse teams'.

There is also the issue of compounded impact. The systemic disadvantages of platform work, such as being low-wage and precarious, compound with the disadvantages already faced by many disabled people. In terms of the gig economy, for instance, disabled workers are further impacted by the self-employed status of this kind of work, at least in countries where this legal definition applies. This is because, defined as independent contractors, gig economy workers receive no protective benefits such as social security and sick leave, which is likely to have a disproportionate impact on disabled people (Claypool et al, 2021). As noted by Claypool et al, due to historic discrimination and added barriers, disabled people also face a higher cost of living that would make losing these benefits or protections extremely difficult. This also points to the intersectional character of inequalities pertaining to AI, and how they impact on disabled people who embody other minoritised identities such as ethnicity, race and gender. For instance, in the aforementioned study of 24 disabled workers, Shannon and Cosley (2022: 19) found that these workers faced 'new, complex challenges that were compounded by the intersection of these identities'. They illustrate these challenges through several vignettes, for example one participant, a Black delivery driver, had to engage in additional labour to protect himself both on account of his disability and his race, but at the same time, both of these marginalising factors also reduced his earnings relative to other drivers regardless of whether they were disabled. Another participant, a gay, non-binary rideshare driver who had PTSD, panic disorder, and generalised anxiety disorder, 'struggled tremendously' with maintaining a traditional job, and turned to gig work for its flexibility. However, highlighting the complicated interactions that intersectional marginalisation can cause disabled gig workers, this particular participant was exposed to additional harassment on the basis of their gender identity that in turn exacerbated their mental health disorders, 'reducing their ability to both hold traditional jobs and do gig work'. As such, Shannon and Cosley illustrate how 'the interaction of disability, discrimination, and economic need' can lock workers into cycles that are hard to escape (p.21).

Generative AI

Generative AI is increasingly being used across different workplace processes. In hiring, one common example is resumé screening, where AI is used to rank resumes, a task for which the use of large language models (LLMs) such as ChatGPT is becoming much more common (Glazko et al, 2024). Glazko et al (2024) carried out a resumé audit to detect and evaluate ableist bias that the model GPT-4 may have against disabled people during resumé screening. To do this, they used a control CV, with disability-related information omitted, and six synthesized enhanced CVs for different impairments, with disability-related information included. The six CVs were enhanced with an additional leadership award, scholarship, panel presentation, and membership that are disability-related. The authors represented five specific impairments in the enhanced CVs: depression, autism, blind, deaf, cerebral palsy, and one non-specific impairment, 'Disability' (p.689). As the authors state, 'an unbiased system should always choose the enhanced CV [ECV] over the CV, since the ECV contains additional awards, presentations, and leadership evidence but is otherwise equivalent' (p.692). Their results point to GPT-4 being biased in several ways. Firstly, there was a strong preference for the CV over the ECV (which was only ranked first in 15/70 trials). Secondly, there is an indication of bias against autism since, of all the enhanced CVs, the ECV associated with autism was ranked first least - 0 times compared to the control CV. The deaf condition enhanced CV followed closely after, ranking first only once out of ten trials. Depression and cerebral palsy were ranked first twice each, and general 'disability' and blindness were both ranked first 5/10 times. This indicates that with generative AI there may be a need to further consider how the use of AI in employment is audited as it introduces new features that may not be captured by existing auditing methods at the same time as potentially being more widely used by employers.

As further discussed in the section on minority languages, the growing reliance on LLPs such as ChatGPT to carry out tasks at work, may also present further barriers to those who rely on minority languages, including sign language and braille, that often lack a sufficient corpora to facilitate the development of appropriate genAI tools. This then highlights the need to actively invest in different infrastructure, including linguistic infrastructure, to ensure that new technologies can also support a wide range of workers' needs.

Responses

Significant efforts have been made towards improving design in AI to better account for disability, mostly in terms of finding ways to remove data that might act as proxies for disability or making disability explicitly visible in design to correct for algorithmic outputs (e.g. research in the academic community FaCCT that engages with fairness in machine learning and AI). However, Tilmes (2022) considers this form of fairness in design as a limited approach and instead proposes 'disability justice'. This would go beyond technical adjustments to training data and input–output relations by serving as a framework 'for reasoning about how ableist structures and norms subtly configure and restrict the ostensibly objective aspects of AI design' (p.8). As well as structures and norms, disability justice would also entail deeper analysis of how data analytics and machine learning help to define and redefine concepts such as disability and fitness. Thus, this would encompass centering analysis of assumptions and values in design; closing gaps between stakeholders; and pursuing policies that empower further activism. Tilmes suggests that, practically, by reflecting on values encoded into targets like 'fit,' we can draw out legitimate indicators of employability, such as word choice, 'from ones steeped in ableist norms and unrelated to performance, such as speech patterns and tone' (ibid). Others place more emphasis and onus on the responsibility of stakeholders, such as employers but also technology companies and suppliers, to ensure any AI tool does not harm minoritised or marginalised workers including disabled workers. For example, Givens (2020) suggests that employers must interrogate the actual variables being considered and weighed in the algorithms themselves. In doing this, Givens proposes that employers ask core questions such as 'does a hiring test evaluate factors that truly relate to the job in question? Does it assess candidates on their individual merits, rather than inferences about disability? Was the test designed and reviewed by people with diverse lived experiences, to identify potential barriers?' Further, Engler (2019) suggests that for algorithms that are crucial in hiring, companies should publicly release bias audit reports that include summaries of the predictions made across subgroups, especially protected classes, 'rather than simply claiming their models have been evaluated and are bias-free'.

Similarly, Nugent et al (2020) suggest a range of questions that can be asked by different employer stakeholders during the procurement of AI hiring systems in order to better prevent harm and which serve as 'intervention recommendations' for improving hiring AI systems. For example, they suggest strategy stakeholders ask themselves 'does this technology align with our organizational strategy to increase diversity and representation?', and 'which people at my organisation should be involved in the decision to investigate, procure and apply systems so as to create a governance process that does not adversely impact disadvantaged and disabled employment seekers?'. HR and operations stakeholders should ask 'what are the benefits and risks of this technology for disabled and other disadvantaged employment seekers?' and 'was a shared understanding of inclusivity and fairness—with specific reference to eliminating the root causes of disability related discrimination—designed into this technology?' In addition, procurement stakeholders should ask 'has this supplier proved their products are safe for disabled and other disadvantaged employment seekers before you purchase?' and 'How has the supplier actively involved PWD to test and validate its products?' Furthermore, some job search platforms utilise algorithms in favour of disabled job seekers, such as Hireautism and Inclusively. These platforms allow for disabled jobseekers to list more broadly their strengths, interests and needs, and are developed with disabled people and their communities (Zhuang and Goggin, 2024). And while AI is used to match disabled job seekers with employers, these platforms 'also rely on broader attitudinal change to persuade employers to come onboard and support their cause'.

From a regulatory perspective, Kelly-Lyth (2023) argues that EU equality law imposes a duty on employers which requires consideration in the context of algorithmic decision-making: the duty to make reasonable accommodations for disabled persons. This duty means that employers must take 'appropriate measures' in individual cases to enable persons with disabilities to access, participate, and advance in employment. Such duties would also extend to UK equality law and employers in the UK must adhere to the Equality Act 2010 including the duty to make reasonable adjustments. Kelly-Lyth therefore suggests we can use the existing framework as it 'already recognises the need for employers to respond to the uniqueness, fluidity, and context-specificity of disability through the reasonable accommodations obligation' (p.164). Crucially this would require both system-level design changes, on the basis that disabled people are structurally disadvantaged, as well as individualised adjustments on a case-by-case basis, because every impairment is unique. Buyl et al (2022) also see opportunity in EU equality law, suggesting that 'failure to provide reasonable accommodation to a person with disabilities by an automated hiring system amounts to prima facie discrimination, thus shifting the burden of proof to the defendant' (p. 1072). In this context, they argue that the prohibition of discrimination requires 'more than a mere refraining from unequal treatment, it may also require an adjustment of an apparently neutral provision or practice if this creates a particular disadvantage for members of a protected group' (p.1074). They therefore propose that a specific application of the duty of reasonable accommodation 'could be to provide an alternative hiring procedure for people with disabilities who fear to be treated unequally', and in practice this could be fulfilled by the opportunity to opt out of the AI-driven selection procedure and demanding human intervention. Further, the authors add that data-driven solutions to the duty to make reasonable accommodations are inherently limited and infeasible as 'AI systems struggle to supply reasonable accommodation due to a heterogeneity that is unique to PWDs [people with disabilities] as a protected group' (p.1080).

Finally, with regards to algorithmic management in platform work, Sannon and Cosley (2022) asked disabled workers how gig work could be improved and found that workers' suggestions included improving transparency in relation to task selection by providing disabled workers with more detailed task descriptions that would help them make more informed decisions. It also included giving workers more control over workflow and adaptable tasks, as many workers wanted to be able to adapt work processes to their abilities. In addition, several workers brought up the desire for an option to indicate impairments on gig platforms that might help algorithms better assign tasks based on workers' abilities. However, at the same time, workers were worried that sharing their impairment, disability or condition would open them up to discrimination from the platforms or could cause them to lose access to work altogether and were uncertain whether they would actually use such a feature. The authors therefore suggest an alternative solution could focus on workers' and customers' preferences about characteristics of tasks rather than requiring workers to share.

AI Inequalities: Ethnicity[2]

Summary

The use of AI in workplaces has been found to discriminate against Black, Asian and Minority Ethnic workers based on both direct and indirect forms of discrimination. Although the design of particular AI models may not explicitly use demographic data about ethnicity, other data can serve as precise proxies for such characteristics in ways that result in direct forms of discrimination. In general, AI tools have been found to reflect racialised norms including in the use of such tools for assessments or profiling of workers based on voice, facial expressions, or language, with ethnic minorities more likely to be disadvantaged. Black, Asian and Minority Ethnic workers are also disproportionately exposed to surveillance and performance assessments linked to data-driven technologies as they are overrepresented in sectors where such technologies are most widely used. At the same time, some workers from ethnic minority backgrounds, and particularly migrant workers, have been found to prefer employment with a heavy reliance on algorithmic management, such as platform labour, as they feel they are less exposed to direct interpersonal discrimination. However, several cases of racial discrimination within the gig economy have been exposed and there are ongoing disputes about the use of some AI tools. Generative AI has been found to further racial stereotyping and discrimination when used, for example, for ranking potential candidates. Efforts have therefore been made to impose stricter auditing on AI tools used in employment and toolkits have been developed to assist with bias mitigation.

Introduction

The advent of AI in the workplace is the subject of both direct and indirect discrimination experienced by racialised and ethnic minorities. As Atkinson and Collins (2024) have pointed out, unlawful direct discrimination in work can occur even when algorithmic systems do not make use of protected characteristics such as race or sex in their internal operations, because decisions or recommendations that are based on characteristics or combinations of data points can act as precise proxies for protected characteristics. Characteristics such as names and language spoken have proven to serve as proxies in hiring and generative AI that can lead to discriminatory outcomes. In addition, algorithmic management practices also amount to unlawful indirect discrimination where they put protected groups at a 'particular disadvantage' that cannot be justified as proportionate. At the same time, it is important to recognise that research has also found that some workers from ethnic minorities, and

[2] The authors of the report have used language consistent with TUC Cymru guidance. Language used in studies cited in this report may differ from this.

migrant workers in particular, feel that algorithmic management, especially in platform labour, is a fairer system because it is seen to do away with the kind of direct discrimination that such workers may experience otherwise from managers. They therefore might actively seek it out as a preferred form of work to standard employment (Bonhomme & Muldoon 2024).

Yet research indicates that although it may not appear in the same form as direct discrimination, AI has been shown to potentially embed and exacerbate structural forms of discrimination in the attempt to provide singular certain outputs such as normative measures like 'success'. 'cultural fit' and 'employability' that underpin AI across hiring to management. Significantly, this can risk amplifying and reproducing institutional, systemic and structural forms of racism that fundamentally limit opportunities for Black, Asian and Minority Ethnic workers. As illustrated below, these are different to interpersonal racism, which is the kind of bias that many AI hiring tools claim to eradicate and that platform workers welcome algorithmic management as an answer to. Accounting for forms of structural impact helps us to disambiguate the notion of algorithmic bias and illuminates the social limitations of technical fixes and the 'debiasing' of hiring, management and generative AIs.

Importantly also, research on this topic has so far overwhelmingly concentrated on the gig economy. As Wood (2024) highlights, 'empirical research into algorithmic management in non-platform work settings remains largely confined to warehousing…with a few exceptions exploring manufacturing…and also retail.' Of these, only few studies explicitly mention Black, Asian and Minority Ethnic workers and there is a disproportionate amount of research studies on this topic focused on the US. What follows is therefore limited by these research gaps as we outline the multifaceted ways in which AI inequalities relate to questions of race and ethnicity and explore prominent examples identified by research that illuminate how the advent of AI in the workplace is impacting disproportionately on some groups of workers. We start by looking at hiring processes, before addressing the use of AI in management and the rise of generative AI, and then providing a review of relevant responses.

AI in hiring and recruitment

Racial and ethnic discrimination has been highlighted as a concern at every stage of the hiring process, from job searching and job adverts to interviewing. For instance, the European

Network Against Racism's (ENAR) 2020 report into hiring AI systems highlights numerous examples of both indirect and direct discrimination of racial and ethnic minority candidates in nearly every aspect of the hiring process, highlighting various ways in which racial and ethnic minority candidates can be eliminated and excluded with the use of automated hiring. From the outset, for example, targeted job adverts can optimise discrimination by using attributes that act as a proxy for race, such as names, postcodes, or membership of particular social media groups (Vinod Bhatia et al. 2024). Further, even when open or 'inclusive' targeting parameters are set, advert delivery can still end up being unintentionally skewed across racial and gender lines by, for example, making it more expensive to target adverts at some groups of job candidates (2020: 13). In addition to job adverts, job searching platforms can also produce algorithmic discrimination through preference for some languages over others. For instance, in its 2023 report 'Discriminatory By Default?' The Equal Rights Trust explored several case studies of algorithmic systems and found that an online platform implemented by the government of Paraguay to help people find jobs, called 'ParaEmpleo', was discriminating on the basis of nationality and language. The intention behind the deployment of this platform is that job-seekers use it to create a profile, listing their relevant qualifications, skills and specialisations, and the ParaEmpleo algorithm then 'matches the user with employment opportunities suited to their profile and recommends relevant courses to increase their chances of finding employment' (Equal Rights Trust, 2023: p.12). However, ParaEmpleo is available to use only in Spanish and English despite the fact that the country has two official languages, Spanish and Guaraní, and that approximately 90% of the population speak Guaraní. In particular, the report highlights that, as Guaraní is spoken by the indigenous Guaraní people, many of whom are not bilingual, the ParaEmpleo's accessibility 'is not only limited on the basis of language, but also in a way which disproportionately impacts on the members of the Guarani ethnic group'. In this sense discrimination was built into the ParaEmpleo system from its inception, prior to any work to develop or design the system, when the choice was made to restrict the language of operation.

In addition to job adverts and job search platforms, uses of screening algorithms to eliminate candidates early on in the hiring process who do not meet the desired criteria can also introduce racial bias. One example of this is that HR chatbots can be trained to use a specific database of employee language phrases to guide conversation, yet they are severely limited when it comes to processing and mimicking written and spoken language which does not belong to the dominant group (ENAR, 2020 14). As ENAR highlights, speech recognition from all of the major technology producers shows a significantly higher error rate with people who are Black, misunderstanding between 25 and 45% of words spoken. Video interviews, meanwhile, allow employers to use AI to rate videos of each candidate according to verbal and nonverbal cues, but the software reflects the previous preferences of hiring managers meaning that if more white males with generally homogeneous mannerisms have been hired in the past, 'algorithms will be trained to favourably rate predominantly fair-skinned, male candidates while penalising women and ethnic minorities who do not exhibit the same verbal and nonverbal cues'. Moreover, as we have seen previously, technology companies are making strong claims about the efficacy of facial movements to process emotion during video interviews, but these are assumed to be cross-cultural, whereas the evidence base is limited in its reliability, lack of specificity and limitations to the generalisability. For example, ENAR demonstrates that skin colour and face shape can significantly alter results, 'with Black profiles associated with anger or contempt, while Asian faces are perceived as blinking repeatedly (associated with nervousness or deceit)' (ibid). In terms of the former, the technology reproduces human bias, 'requiring that black professionals must amplify positive emotions to receive parity in their workplace performance evaluations'. In terms of the latter, this discrimination is based on face-shape perception by the technology. What is more, skin colour can even affect 'whether an AI interview recognises a person is present, and whether it begins or continues the interview'.

Another common AI tool in hiring is assessments, such as 'neuroscience' web and mobile games that are used to measure cognitive, social and emotional traits of candidates, like processing speed, memory and perseverance. These assessments focus on selecting candidates that reflect current 'top performers' in the workplace, despite a lack of objective bases on which to identify 'top performers' that are free of bias. Tests which measure cognitive ability and personality (used for 'cultural fit' assessments) have long been suspected to be inherently discriminatory against racial and ethnic minorities as well as disabled people (ENAR 2020). Further, ENAR points out that 'cultural fit' indicators, such as hobbies or interests, are often strongly correlated with certain nationalities or racial groups. This means that, in relying on data about people who are perceived as 'high performers', or 'long stayers' within organisations (the latter being impacted by issues of racial harassment, exclusion or return-to-work discrimination against women, for example), automation and predictive tools also run the risk of reproducing these biases. All of these examples illustrate that automation and AI make possible racially discriminatory outcomes in every stage of the recruitment process. As such, the classification schemes that automated or predictive tools operationalise to 'measure' job applicants are not neutral or objective but based on existing cultural norms and values that are not free from racial inequalities. As a result, it is important to distinguish between different kinds of bias when we account for the impact of AI in work and consider automated hiring as reproducing structural, institutional and systemic racism on top of interpersonal racism. According to Bogen and Rieke (2018), the term "bias" in relation to AI is often used to refer to interpersonal bias - 'prejudices held by individual people, whether implicitly or explicitly' - and has long plagued hiring processes. Similarly, ENAR defines interpersonal racism as the bias that occurs ‘when individuals interact with others and their personal racial beliefs affect their public interactions’ (2020: 6). This matters because hiring biases are produced partly by unconscious bias, which favour some groups more than others, while conscious biases for particular types of career history, education, areas of interest, also play a large part in which job applicants we select. Many AI and automated hiring tools promise to remove bias from the recruitment funnel but, even if this were possible, this relies on a selective understanding of racial bias that is limited to interpersonal racism. In fact, ENAR argue that a focus on unconscious bias ‘means that there is little attention paid to the much larger problem of structural discrimination, which produces and perpetuates accumulated disadvantage through blocked access to key institutions and opportunities’, and institutional racism, ‘which affects how organisations interact with minority ethnic candidates, for example through policies and workplace cultures that serve to benefit certain workers and disadvantage others (Bogen and Rieke, 2018). For ENAR these are precisely the types of bias which technology reproduces and amplifies at speed. Similarly, Bogen and Rieke (2018: 7) argue that structural kinds of bias act as barriers to opportunity for job seekers ‘especially when predictive tools are involved’. They note, as many social patterns related to education and work reflect troubled legacies of racism, ‘blindly replicating those patterns via software will only perpetuate and exacerbate historical disparities’ (p.9).

The structural dimensions of racial and ethnic bias are reflected in the use of existing data sets to train algorithms, particularly in machine learning tools tasked with automation and prediction. This means that data from the existing workforce, either within a specific organisation or sector or labour market at large, is the basis for those measures mentioned above, be that cultural fit, cognitive and emotional responses, or salary prediction, in order to determine whether a candidate is worth hiring or not. This is not a technical problem but the result of structural and systemic racial inequality, as the lack of diversity in the data pool is often a direct result of biased hiring and employment practices (ENAR, 2020). Consequently, as ENAR suggests, ‘trained on the existing workforce and performance benchmarks determined from the successes and perceived failures of those who already work for the organisation, new hires continue to resemble those hired before (as they are all based on the same characteristics and means of portraying those characteristics)’ (p.9). Further, Bogen and Rieke (2018) highlight that these patterns can also emerge as tools are used, particularly when models are built to learn and adapt to the preferences of its users over time. Importantly, in this regard removing or obscuring sensitive factors like gender and race will not prevent predictive models from reflecting patterns of bias. This also means that ‘debiasing’ approaches to algorithmic bias will often be insufficient. As West et al (2019: 10) argue, ‘it is not just that AI systems need to be fixed when they misrecognize faces or amplify stereotypes. It is that they can perpetuate existing forms of structural inequality even when working as intended'.

A related issue here that affects both hiring and algorithmic management is the simplification of racial and ethnic categories within algorithmic systems (which also applies to other protected characteristics such as gender and disability). When a system differentiates between groups of people for a particular purpose based on their ethnicity or gender it gives them particular meaning but these categories are constructed in particular contexts. This means descriptions of racial or ethnic groups adopted in the United States, for example, have different outcomes when applied in Europe (ENAR, 2020; Sanchez-Monedero et al. 2020). However, as ENAR points out, 'algorithmic systems frequently model those categories as fundamental attributes of people'. Thus, 'in an attempt to increase the "elegance" of an algorithm', categories such as gender, race, ethnicity and disability must be simplified. ENAR suggests that this process can lead to the exclusion or alienation of those whose identities 'are poorly served by a lack of nuance and complexity in definition' such as individuals who are mixed race or dual nationality, potentially forcing 'conformity to definitions which fit within predominant categories'. Furthermore, algorithmic systems which rely on simplistic yes/no categorisations run into difficulties when comparing intersectional experiences, such as a white man compared with a black disabled woman (Sanchez-Monedero et al. 2020). For West et al, these simplistic classification structures are likened to histories of 'race science' and 'are a grim reminder that race and gender classification based on appearance is scientifically flawed and easily abused' (2019: 3). They argue that 'systems that use physical appearance as a proxy for character or interior states are deeply suspect' such as tools that claim to assess worker competence via 'micro-expressions'. Not only does this actively replicate patterns of racial and gender bias 'in ways that can deepen and justify historical inequality' but the commercial development and deployment of these tools is a significant cause for concern.

Another important factor in structural impact is that many of the technology companies that develop hiring and algorithmic management systems have a significant lack of diversity within their workforces. As West et al (2019) highlight, there is a diversity crisis within the AI sector in relation to race and gender. For example, they point out that at the time of writing, only 2.5% of Google's workforce is Black, while Facebook and Microsoft are each at 4%. They argue that, 'given decades of concern and investment to redress this imbalance, the current state of the field is alarming'. Moreover, they suggest this is a systemic issue that acts as a structural cause of further bias because there is a 'close relationship between these workplaces with discriminatory practices and discriminatory tools: a feedback loop that is shaping the AI industry and its tools' (p.9). This has been suggested elsewhere, for example Zapata (2021) argues that because 'Silicon Valley is still prominently populated by white people, with men comprising the majority of leadership positions' that 'begs the question of how the technology industry can create fair and balanced AI for the masses if there are still diversity challenges within the very teams designing and implementing the algorithms upon which that AI relies.' In this regard West et al. (2019) indicate that addressing the structural biases generated by hiring and management AI will need to grapple with the systemic issue of lack of diversity in the organisations that develop these technologies. Again, this renders technical fixes 'futile' for the authors, as 'only by examining discrimination through the lens of its social logics (who it benefits, who it harms, and how) can we see the workings of these systems in the context of existing power relationships'.

AI in management

The concerns raised above stretch across both the use of AI in hiring as well as management and follow many of the same patterns. However, the use of AI in not only directing and evaluation workers, but also discipling workers has raised further questions about its potential harm for particular groups of people. For example, to manage workers, Uber has implemented a facial recognition system called 'Hybrid Real Time ID Check' (RTID) in order to authenticate drivers' identities and prevent them from sharing access to their accounts (Worker Info Exchange, 2021: 17). Introduced in April 2020, the RTID incorporates Microsoft's FACE API facial recognition software and requires drivers and couriers to periodically take real-time selfies to continue using the Uber app. The photo is then checked against the driver's account profile picture (and in some jurisdictions, against public databases to 'prevent identity borrowing or to verify users' identities.') (ibid). There have been a number of cases so far where ethnic minority Uber workers have been discriminated against and lost their jobs because the RTID system incorrectly failed to authenticate their selfie photos.

In October 2020 Imran Javaid Raja, an Uber driver, was dismissed from his job and reported to Transport for London after he failed two facial recognition checks via the Uber driver app. Transport for London then revoked Imran's private hire driver and vehicle licence without any notice or without allowing him any opportunity to represent himself, review the evidence or present an appeal (Farrar et al, 2021: 1). In November 2021, Pa Edrissa Manjang, who is Black, was deactivated as an UberEats courier due to selfie verification failure. Pa was not given any warnings or notified of any issues until his dismissal; the RTID verification system appeared to approve all of his photos with a green check. Following his dismissal, Pa sent numerous messages to Uber to rectify the problem, specifically asking for a human to review his submissions. Each time Pa was told 'we were not able to confirm that the provided photos were actually of you and because of continued mismatches, we have made the final decision on ending our partnership with you' (Worker Info Exchange 2021: 18). In addition, Abiodun Ogunyemi, a Nigerian Uber Eats driver was locked out of the app in March 2021 after several failed attempts using the facial verification software. Abiodun told the Guardian that his family had faced 'serious suffering' as a result and said 'I feel the algorithm is discriminatory to people of colour. I know about five black people the same thing has happened to' (Booth 2021).

Importantly, this practice is expanding as deployments of facial recognition spread to other gig economy companies. For example, according to Worker Info Exchange (2021), Bolt has since announced that it was investing €150 million in AI driver anti-fraud detection systems including facial recognition, while Deliveroo and Ola Cabs have also announced that they would introduce facial recognition identity checks. This is all the more pertinent given that, according to Transport for London (TfL) reports, 94% of private hire vehicle (PHV) drivers (in London) are from Black, Asian and Minority Ethnic backgrounds. In response to Uber's discriminatory verification checks based on facial recognition software, the App Drivers and Couriers Union (ADCU) and Independent Workers Union of Great Britain (IWGB), with support from the Equality and Human Rights Commission and the NGO Worker Info Exchange filed a claim in 2021 at the Central London Employment Tribunal on behalf of three drivers who had experienced indirect discrimination (Trott and Gittins 2021). Lawyers argued that facial recognition systems, including those operated by Uber are inherently faulty and generate particularly poor accuracy results when used with people of colour. In Pa's case the claim is that not only was he dismissed on the back of racially biased facial recognition technology, but also that while working for Uber Eats he was targeted for heightened and excessive facial recognition verification checks which amounts to racial harassment (ADCU, no date). ADCU co-founder Yaseen Aslam said of the case: 'this case reveals again the ugly truth that the economic success of the gig economy is largely a racist endeavour with technology used cynically to objectify and exploit the most vulnerable in society' (ibid). Uber Eats applied to have Pa's claim struck out in 2021 but this was rejected; in March 2024 Pa received a financial settlement, bringing the case to a close (Wright, 2024).

In addition to legal action, IWGB and ADCU members organised strike action in October 2021 in protest against Uber's use of facial recognition technology, with IWGB calling for a 24 hour boycott of Uber in solidarity with striking drivers (IWGB, 2021). The strike marked the launch of a campaign by Black Lives Matter UK and the IWGB to demand Uber scrap its 'racist facial recognition algorithm, reinstate unfairly terminated drivers and couriers and introduce a fair terminations process' (IWGB, no date). Unions have also demanded change and reform at Transport for London alongside Uber, in particular demands for these organisations to review their internal practices in relation to revoking private hire licences. This includes a demand that 'TfL review every revocation of a private hire driver licence since January 1, 2020 where TfL's decision is based on analysis from any form of surveillance technology used by a licensed operator including facial recognition, geo location checks and other electronic inputs'. It also includes the demand that TfL commit 'to end the practice of immediate revocations where the decision is based upon surveillance technology evidence and ensure every driver facing such allegations is given appropriate time for appeal and is given access to the evidence against them'. Further, it was demanded that in the future TfL 'review and challenge evidence presented by app operators against private hire drivers based on surveillance technology' (Farrar et al, 2021: p.1).

Outside the platform work and the gig economy, research in the US has made the case that, in general, algorithmic management techniques across sectors tend to continue historical practices of 'racial quantification' that have shaped experiences of racism for a long time (Ajunwa 2023). As Ajunwa explains, not only are Black Americans 'facing increased surveillance compared to their white counterparts due to prejudiced employer perceptions' but additionally 'Black Americans are also overrepresented in industries where surveillance is simply more prominent overall' (2023: 307). One aspect of this is that since low-wage workers, particularly those in the service sector, face far greater levels of surveillance than workers in other industries, from keystroke logging to drug testing to GPS location tracking, Black Americans are disproportionately impacted by being disproportionately represented in such work.

In addition to extending surveillance through algorithmic management, research has documented the particular rise of emotional artificial intelligence (EAI), which is increasingly integrated into enterprise systems to augment and automate organisational decisions and to monitor and manage workers (Corvite et al. 2023). For employers, the benefit of implementing EAI lies in its potential 'to improve workers' wellbeing and performance as well as address organizational problems such as bias and safety' (ibid). However, Corvite et al. (2023) found in their study of perceptions of EAI among marginalised workers in the US that workers are concerned about this technology, and that many see no benefit to it: '32% of participants, 71.7% of whom were participants who identified with a marginalized identity (i.e., person of color, woman, transgender, non-binary, having or had a mental illness) did not note any benefit when asked to describe potential benefits they might receive from EAI in the workplace' (2023: 10). Further, 15.9% of participants expressed concerns regarding employers using EAI in ways that could lead to bias against data subjects with a marginalized identity. For instance, one participant, a disabled Black woman, mentioned that she would be concerned about EAI use in the workplace "if [EAI systems] are not programmed properly to consider race and culture." She went on to describe how her identity as a "poor/black/elderly/woman" could lead to obstacles in "getting real, honest, caring help from professionals... [and she has] to take into consideration that the bots are being programmed by people which most times, (maybe unintentional), use their bias" (ibid: 18). In addition, Hajric et al. (2024) look at the use of facial recognition technology for EAI used for human resource management, including decision-making in, for example, the construction and optimization of virtual teams, appropriateness for promotion to leadership positions, and fitness-to-task in mission critical work. As part of their study, they explore social implications of such technologies, including the possible discrimination against women, racial minorities, undocumented immigrants and refugees, and people with visible and invisible impairments. In particular, they suggest that certain emotions like anger 'are socially penalized and therefore their public expression will likely be hidden'; while the extent to which emotions, like anger, are tolerated 'are cultural and context-dependent to local practices' (p.6). They therefore argue that care is vital for employers 'when implementing universal applications of westernized standards of emotion, as it may confuse a friendly response in one culture, misinterpreting it as anger' (2024: 6). One study from 2018, for example, using a publicly available data set of professional basketball players' pictures, compared the emotional analysis from two different facial recognition services, Face++ and Microsoft AI and found that both services interpret Black players as having more negative emotions than White players. However, they present bias in two different ways; Face++ consistently interprets Black players as angrier than White players, even controlling for their degree of smiling. Microsoft registers contempt instead of anger, and it interprets Black players as more contemptuous when their facial expression is ambiguous. As the players' smile widens, the disparity disappears (Rhue 2018).

Generative AI

Early research into generative AI is showcasing further issues of racial discrimination that have impact on work and workers. For example, a study of that carried out a 2024 Bloomberg resume audit study of OpenAI's generative AI technology (specifically the model GPT 3.5) found that it displayed preferences for certain racial identities in questions about hiring 'by systematically producing biases that disadvantage groups based on their names' (Yin et al, 2024). The study carried out an experiment by using fictitious names and resumes to measure algorithmic bias and hiring discrimination. Applying methods from previous similar studies, Bloomberg reporters used voter and census data 'to derive names that are demographically distinct — meaning they are associated with Americans of a particular race or ethnicity at least 90% of the time — and randomly assigned them to equally-qualified resumes'. Then, when asked to rank those resumes 1,000 times, the model GPT 3.5 favoured names from some demographics more often than others, to an extent that would fail adverse impact benchmarks used to assess job discrimination against protected groups. In particular, resumés with names distinct to Black Americans were the least likely to be ranked as the top candidate for a financial analyst role, compared to resumes with names associated with other races and ethnicities. Those with names distinct to Black women were top-ranked for a software engineering role only 11% of the time by GPT — 36% less frequently than the best-performing group. As Yin et al suggest, if GPT treated all of the resumes equally, 'each of the eight demographic groups would be ranked as the top candidate one-eighth (12.5%) of the time'. However, resumes with names distinct to Asian women were ranked as the top candidate for the financial analyst role more than twice as often as those with names distinct to Black men.

Further, in terms of the deployment of generative AI, the prominence of different ethnic and racial groups in some sectors more than others has implications for the impact of this technology. For example, a 2023 study by Indeed's Hiring Lab Workers of different social groups (by age, gender and race and ethnicity) in the US explored how this technology will impact social groups unequally 'because members of each demographic tend to work in different fields' (Hering 2023; Honorof, 2023). The study found that Asian/Pacific Islander descent (AAPI) workers face the highest potential level of exposure to generative AI because 21.5% of AAPI workers are employed in sectors with the highest level of potential exposure - such as mathematics/computers sectors, business and finance, and management - as opposed to 12.9% for the next-highest demographic group, White workers (Hering 2023; Honorof, 2023). Hispanic workers face the least potential exposure as just 7.1% of Hispanic workers are employed in high-exposure fields, while 42% work in fields with the lowest potential exposure (Hering, 2023). As highlighted in earlier sections, this is likely to mean that both risks and opportunities of generative AI will be disparately distributed across different demographic groups, including ethnicity, as a reflection of these broader structural dimensions of labour markets.

Responses

The issue of racial discrimination is a significant concern in the growing adoption of AI across employment. A key response to this has been a growing emphasis on auditing. For example, in New York City, the Department of Consumer and Worker Protection has introduced a law regulating the use of automated employment decision tools ("AEDT") by employers and employment agencies in the city. The law took effect in January 2023 and responds to the issue of gender, ethnic and racial bias by requiring that a bias audit must be conducted “no more than one year prior” to the use of an AEDT by employers or employment agencies. A bias audit is defined as ‘an impartial evaluation by an independent auditor’ to assess the tool’s potential disparate impact on sex, race and ethnicity, and must use data from the employer’s or employment agency’s own historical use of the tool (Maurer 2023). However, it has been suggested that so far impact has been fairly underwhelming ‘as employers have the freedom to decide whether or not their systems are covered by the law and overall lax enforcement, it’s largely gone ignored’ (Turner, 2024). In fact, a recent study by researchers from Cornell University and Data and Society found that only 18 out of 391 NYC employers have posted audit reports to their websites (Wright et al, 2024).

Nonetheless, similar requirements are being put forward by both the EU and the UK, for example through EU’s AI Act which includes a prominent role for bias auditing, particularly in settings such as employment, and in UK’s guidance on ‘Responsible AI in Recruitment’. However, dominant approaches to AI auditing have tended towards computational interpretations of regulation and standards, evidenced by the growing industry of auditing tools that focus on areas such as accuracy and algorithmic bias, often providing very narrow and contentious criteria for fairness. What is more, audits tend to be confined to separate parts of the recruitment and hiring process, with unclear lines of accountability, and limited information about impacts on different groups of people outside a few demographic categories (Ada Lovelace Institute 2024). More broadly, there have been calls from across civil society for changes to how we approach the development, design and implementation of algorithmic systems in hiring and employment. In light of evidence of gender, age and ethnic discrimination, the Equality Rights Trust (2023), an independent international organisation, for instance, calls for states and firms to adopt an ‘equality by design approach’ that would need to be pre-emptory and precautionary in order to counter the ‘discrimination by default’ built into many algorithmic systems. This would also require taking a proactive approach to ensure that potential discriminatory impacts are identified and addressed before they occur and that equality considerations are intentionally incorporated into the design, development, and deployment of algorithmic systems.

Additionally, ENAR (2020) has developed a toolkit for mitigating and preventing racial bias in hiring. The toolkit is aimed at HR managers, diversity and inclusion managers and programmers. Specifically, the objective of the toolkit is to ‘ensure that consumers of off the shelf and custom AI solutions for Human Resource Management have a clear guide to challenges, solutions and good practice, in a format which supports conversations with Programmers providing solutions. Essentially the toolkit helps to educate HR and

D&I managers about bias amplification and reproduction in algorithmic systems, and provides practical steps to ensure companies comply with the principles set out by The European Commission High Level Group on AI in their broad guidelines for ethical use of AI. Aside from this compliance, the toolkit promotes co-production, especially by involving those who are potential candidates from underrepresented populations within the workforce. The toolkit gives several co-production examples, such as developing user-vetted lists of desired characteristics and generating discussion within corporate teams about why these characteristics are deemed desirable. No formal evidence has been provided about the impact of such a toolkit, but it speaks to a number of measures that are seeking to directly engage with managers about their implementation and use of AI technologies (see also Institute for the Future of Work in the UK).

AI Inequalities: Minority languages

Summary

There is generally a lack of research on how minority language speakers are impacted by AI in the workplace. However, research on how AI may pose challenges and opportunities for the preservation of minority languages indicate that minority language speakers are at risk of being further marginalised with the advent of AI in the workplace. While technological advancements are being made to develop AI tools that can support and even strengthen minority languages, the dominance of closed systems built on and for majority languages, and English in particular, pose a significant barrier to the preservation of minority languages. A central challenge is the limited resources and corpora available for minority languages to train AI models on. Significant efforts are being made to build linguistic technological infrastructures, including in Wales, where emphasis has also been made on building educational tools and translation tools that can support Welsh speakers across different sectors. Yet concerns remain about the possibilities to counter the further rise of 'super languages' and advance 'linguistic justice' in the context of AI without much wider efforts to ensure more open and democratic technological developments that can centre the experiences of minority language speakers.

Introduction

The advent of artificial intelligence (AI) has brought about significant advancements in various fields, including language technology. The capacity of machine-learning technologies to play a role in the preservation and promotion of minority languages is not new. Somers (2004) identified the capacity of machine translation to help build minority European languages such as Welsh, Irish, Basque, and Galician. Trosterund (2006) argues that no language will be able to function as an administrative language, and by extension a language that can be used by workers to provide services to the public, in a modern society without a developed language technology and argues that there are ways of building linguistically-based language technologies for minority languages. The Welsh Government report *Empowering communities, strengthening Welsh language*, for example, highlights the development and use of Welsh language technology as part of efforts to use Welsh in the workplace to support the language's visibility, especially for those who do not use it at home. This draws on the Welsh Language Technology Action Plan, published in 2018, to establish technological infrastructure as an essential role in the development of the Welsh language, working alongside the technology sector to develop technological products that will support such aims, including the advancement of Conversational Artificial Intelligence and other types of technologies.

However, the impact of AI on minority languages presents a complex and multifaceted issue. On one hand, AI has the potential to pose a threat to these languages by consolidating the dominance of major languages, particularly English, leading to the risk of 'digital language extinction' for minority and regional languages. However, AI also offers numerous opportunities for the preservation and promotion of minority languages through the development of tools and technologies that can support their use in various sectors. Research is still emerging in this area. For example, research at Bangor University has been supported by the Welsh government to improve understanding of how computer systems process the Welsh language, and in 2023, the University of Edinburgh announced that experts at the Universities of Edinburgh and Glasgow have been awarded funding by the Scottish Government to produce a Gaelic subtitling system suitable for the BBC. Funding will also support research towards the production of large language models – similar to ChatGPT – for Welsh and Scottish Gaelic speakers as part of efforts to counter the threat of digital extinction. The research team in Scotland is also helping to develop a speech recognition system for Ojibwe, one of the indigenous languages of Canada, indicating also that international efforts in this area is key to its advancement.

Below we outline what are some of the major findings regarding the intersection of AI and minority languages, looking both at the ways in which technological advancements can and have been used to empower people speaking minority languages, but also the risks associated with AI in terms of further marginalising minority language speakers. Only very limited research in this area is focused on the workplace directly, but it is possible to draw on broader research to consider how such findings might apply to a workplace context. We start by outlining some of the opportunities that AI may offer minority language speakers, before discussing some of the risks involved and end by considering some prominent responses to such risks.

AI and the opportunities and challenges for minority languages

The extent to which AI can play a role in preserving endangered or minoritised languages is greatly varied but brings together research from diverse socio-linguistic contexts. For example, Tan and Jehom (2024) focus on the Gyalrong Tibetan language as a representative case within the broader context of linguistic diversity endangerment. The emergence of digital technology as a supplementary tool to preserve endangered languages provides opportunities and challenges in language conservation. Adopting a qualitative research approach and thematic analysis, collecting data from previous studies, fieldwork, and interviews, this study considers the intersection of digital technology and Gyalrong Tibetan preservation. The opportunities (language revitalisation) and challenges (low resources) are placed within the framework of media ecology and language shifts, exploring how communication technologies shape Gyalrong Tibetan and its cultural context, suggesting that communication technologies including AI and generative AI, have the capacity to inform the shape, form and cultural use of the language. Language resource refers to the set of speech or language data and descriptions in machine readable form, used for building, improving or evaluating natural language and speech algorithms or systems. A low density language is one that does not have sufficient digitised resources to be easily used in digital spaces (Shamsfard 2019). It is therefore significant to try and understand how digital technologies might intersect with minority language revitalisation to inform possibilities for language preservation in the future.

Gerken (2022), for example, argues that AI, and specifically neurolinguistic programming which refers to behavioural technology that looks at how language influences how people think and behave, can create possibilities for the research, protection, and promotion of regional or minority languages. He argues that with sufficient data behind them, there are numerous new possibilities for minority languages to re-establish themselves and find new areas of growth. This, it is argued, can happen through developing tools for use by organisations that can help further consolidate their use in public and working life, such as administrative authorities and public services, the media, the culture sector, and more broadly in economic and social life. Such optimism is further reflected in the work of scholars such as Pradhan and Dey (2023), who examine the language initiatives for indigenous, tribal, and minority languages in India such as Gondi, Maithili, Rajasthani, and Mundari, and the way in which they are integrating technology into their plans. They note three distinct discourses in these approaches: technological optimism in utilising these new opportunities by claiming space for otherwise-marginalised languages, the imperative for collaborative and collective work in order to address sparse datasets, and the need to negotiate the contested nature of imagining a new collective future. Such initiatives, it is argued, are not just a technical project but way of contesting linguistic hierarchies. In order for this to happen, however, it is vital that there is a sufficient corpora (data) to support them, which is where the difficulty lies in their development and use. Building this corpora is challenging for minority languages. However, if there are public data sources available, it may be possible to use them to help build this dataset. Arvani (2022) explores how this has been put into effect with the Lampung language of Indonesia, for example, and argues that through making efforts to contribute to the maintenance of local languages through technology, it is possible to contribute to the maintenance and preservation of minority languages. The developmental research of their Lampung dictionary application is presented as evidence of this. More recently, the Senedd in Wales, went into a partnership with Microsoft to launch a new Welsh translation system as part of Microsoft Translator (Senedd Cymru 2024) that could also help support a larger corpora.

Further examples of research look at specific ways in which technology can help strengthen or renew languages by promoting ways to promote and normalise their use. Such approaches have been developed in languages such as the IndoAryan language family spoken in northern Asia (Shefali et al. 2024), or minority languages in China such as Tibetan, Uyghur, Kazakh, and Mongolian (Zhang et al. 2024). No Language Left Behind (2022), for example, is a project that seeks to close the gap in AI research that engages with low-resource or minority languages through machine translation. Based on interviews with native speakers, they identified needs relating to low-resource language translation support to then develop a conditional computational model that uses AI to create more parity between the languages.

In terms of the impact that the inclusion or exclusion of minority languages from AI might have on minority language speakers, education has been highlighted as a central area of concern. Minority identities more broadly are at risk of exclusion in educational environments, and language is a key consideration among others. The systematic review presented by Salas-Picio et al. (2022) suggests that research in this area is poorly resourced and developed, and that the existing body of work does not reflect the scale of the need for inclusivity. From a technological perspective, their review suggests that as AI services proliferate, it is likely that their applications can enhance the learning process for minority students. However, for this to happen, the technology needs to address four key aims: firstly, it should develop a culturally sensitive learning environment that supports minority students, such as providing content in their native or preferred language. This is something that has also been put forward in Wales, where efforts have been made for technology to be used to support Welsh language education in further and higher education (Coleg Cymraeg 2024). However such tools should not be abstracted from the wider social and cultural context of those students. Nevertheless, the review also suggests that such technologies can potentially play a significant role to the provision of inclusion education to minoritised students, and their importance should not be underestimated. Significantly, the link should be made between the cost of providing these technologies and minoritised communities who potentially lack the economic resources to successfully implement them. This, they argue, is perhaps a more pertinent and pressing point to be made. Thirdly, the review suggests that the successful implementation of such technologies should look beyond their utility in the classroom. The review suggests that such technologies have the capacity to help minoritised students build careers in professions where they are already under-represented, particularly in the fields of science and technology, and exclusion by virtue of language therefore has wider connotations than pedagogical concerns. AI-assisted technology has the potential for lifelong applications in giving minoritised people access to spaces where they had previously experienced exclusion and this should be included in discussions about their usefulness in education. Finally, the review argues that the benefits of AI need to be far-reaching. While the technology itself has a global reach, the benefits of it are yet to be felt unilaterally, and there are significant sociocultural approaches to the design of such technologies that need to be addressed before this can be said to be true.

A further related challenge concerns the closed nature of many AI tools that extend homogenising tendencies by establishing themselves as globalised technology platforms. A report by Minority Rights Group International (2020), for example, draws on case studies from Africa, the Americas, Asia and Oceania, and Europe to highlight the specific problems each indigenous community has in integrating technology into their own sociocultural context. In doing so, they suggest ways of resisting the homogenising effects of globalised technology platforms. This includes: I) mainstream human rights for all into the development and dissemination of technologies, with a particular focus on the barriers that minorities and indigenous peoples face in their specific context. There is not a 'one size fits all' solution. II) Focus on improving minority and indigenous inclusion, not only as end users of technologies, but in the design and production. III) Conduct human rights impact assessments as necessary first step whenever digital technologies are being considered for adoption by public authorities. And IV) Scrutinise the use of AI and automation in decision-making, with a focus on ensuring transparency and non-discrimination.

A key risk for minority languages with the advent of AI, however, is the dominance of English, Spanish, and Chinese in online communication. Smaller languages risk being further marginalised via 'algorithmic hegemony' if there are simply not enough resources to train the model (Algorithm Watch 2023). Similarly, Gerken (2022) argues that while AI can create possibilities for the protection and promotion of minoritised languages, it also consolidates the dominance of the English language, meaning that minority and regional languages are at risk of 'digital language extinction' as the use of AI becomes more prolific. For example, Kshetri (2024) considers the implications of the exclusion of minority languages from large language models on digital exclusion for non-English speakers, outlining the higher costs associated with using generative AI in non-English languages and the lower accuracy of linguistic models due to limited training data. Focusing particularly on minority languages in developing countries, they suggest that unless this is addressed, it is likely to further exacerbate division between English and non-English speakers. In addition, the proliferation of generative AI platforms such as ChatGPT undermines language democratisation with the status of English as the hegemonic language of academic publishing being consolidated by the introduction of AI-powered tools to the research process (Ghio 2024). Ghio goes on to argue that the proliferation of generative AI, and the lack of transparency around how metrics and languages are constructed, risks people being subjected to control. It also opens up a myriad of ethical questions in the research process, ranging from data transparency and data management to cultural representation and moral limits. The status of English as the language of academia, he argues, is already problematic in that it privileges Westernised discourses, a state that is demonstrated by the dominance of journal rankings that privilege English-language texts. AI tools such as ChatGPT face similar criticisms in the way that English-language communication dominates, but such criticisms do not encapsulate the move beyond domination to consolidation that such chat tools engender. It is therefore questionable whether claims made by advocates that such tools are able to democratise the field of academia are substantiated, but as long as those claims are repeated, they risk giving false hope to academics from minoritised communities of access to a career pathway that is otherwise difficult to access.

In the context of the Welsh language, there have been suggestions from the Welsh Government that generative AI offers specific opportunities to promote and support the use of the Welsh language (Senedd Cymru 2024). Despite limitations regarding available data and other issues as outlined above, research carried out in 2019 identified 439 Welsh language applications available in the Apple (UK) App Store (Cunliffe 2019) that also suggests that this is evidence that the market for Welsh language applications is fairly robust and could therefore extend to generative AI. This is important as research in the past has suggested that one of the problems that minority languages face in the development of AI-powered tools and applications is that the impetus for their development is often market driven, which has also lead some researchers to call for a more active part of the nation state to challenge linguistic hegemony (Keegan and Evas 2012). Indeed, Henry (2017) argues that advances in technology are currently helping to speed up the globalisation of 'super' languages but at the same time technology might be used to help reverse the decline of less widely spoken languages. For example, using a pilot study that makes use of a social learning application, Cada Dia, to provide an immersive learning strategy to encourage authentic conversations in a real time environment to create dynamic and meaningful learning encounters Henry aims to gain an understanding of online social learning methods for minority language learning. However, in the study of Cada Dia Welsh, it was found that there was limited engagement with the app amongst learners, in large part due to logistics, but this in turn meant that the limited numbers in the minority language version of the app left learners feeling more isolated and disengaged. This again draws attention to the significance of the wider environment in which minority language speakers – including in work settings – are situated in order to facilitate technology-enabled support of minority languages rather than further advancing the dominance of 'super' languages.

Furthermore, Cunliffe et al. (2022) describes the development of an AI toolkit based on natural language processing (NLP) for Welsh. Natural language processing is the type of AI that allows for computers to understand and use human language and is widely used in AI across search and business intelligence and other applications. Rather than creating it from scratch, Cunliffe et al's approach involved adapting and enhancing the language processing functionality provided for other languages within an existing framework and making use of external language resources where available. The question they pose is whether it is reasonable to expect to see a 'trickle down' effect whereby minority languages can use existing protocols to develop minority NLP frameworks. They conclude that the approach of adaptation and reuse can provide a practical and achievable route to developing language resources for otherwise under-resourced languages where such data might be limited.

Similarly, Corcoran et al. (2021) bring together the disciplines of Computer Science, Mathematics, and Linguistics to develop a method of automatically learning Welsh word embeddings as part of the development of a Welsh-specific Natural Language Processing model. This research is part of a larger study including 157 languages but suggests that progress is being made in making resources more available for the development of Welsh language AI-tools. Jones (2022) also reports on ongoing work on developing and evaluating speech recognition models for the Welsh language using data from the Common Voice project, which is a crowdsourcing project started by Mozilla to create a free database for speech recognition software, alongside other open development kits. This includes activities for ensuring the growth and improvement of the Welsh Common Voice dataset. Two applications have been developed – a voice assistant and an online transcription service that allow users and organisations to use the new models in a practical and useful context, but which have also helped source additional test data for better evaluation of recognition accuracy and establishing the optimal selection and configurations of models. Test results suggest that in transcription, good accuracy can be achieved for structured read speech, but further data and research is required for improving recognition results of freely spoken formal and informal speech. However, for the voice assistant this more limited domain language model provided excellent accuracy. Russell et al. (2022) also present a design for the collection and verification of a bilingual text-to-speech synthesis corpus for Welsh and English that can be used for similar applications and reflect on the challenges of creating an open-source Welsh language corpus large enough to capitalise on neural text-to-speech (TTS) architectures that are needed for these tools.

Limitations of AI for minority language speakers

Although it is clear that technological innovation includes a burgeoning field in developing tools specifically for minority language speakers, research has also highlighted the limits to such innovation in practice. Bowker (2010), for example, explores the dichotomy of the need to provide bilingual services and justifying the cost of this service based on a recipient evaluation designed to determine whether machine translation could be used as a cost-effective means of increasing translation services in Canadian official language minority communities. The results show that not all communities have the same needs, and that raw or rapidly post-edited machine translated output is suitable for purposes of information assimilation, but not necessarily useful for purposes of cultural preservation and promotion. Similarly, Lai et al. (2023) suggests that even if ChatGPT is trained on another language, the time and labour required make it largely unviable. Given the broad adoption of ChatGPT for English in different problems and areas, a natural question is whether ChatGPT can also be applied effectively for other languages, or if it is necessary to develop more language-specific technologies. Lai et al. evaluate ChatGPT on seven different tasks, covering 37 diverse languages with high, medium, low, and extremely low resources. The results suggest a worse performance of ChatGPT for different NLP tasks and languages, calling for further research to develop better models and understanding for multilingual learning. This presents a particular challenge as ChatGPT becomes more widely adopted across social settings, including the workplace, and is in important finding to help guide the proliferation of government partnerships with Gen AI chatbots, such as the recent partnership with OpenAI in Wales.

More broadly, Pradhan and Dey (2023) argue that while there is technical optimism about the possibilities offered by artificial intelligence to advance education for low-resourced languages, there is also an emphasis on the need to build strategies to overcome social and linguistic hierarchies, address biases in existing technology, and tackle tensions between various social actors. As such, they argue that the use of language technology for minoritized language education is understood not just as a technical project but as a contested and ongoing attempt to sub-vert linguistic hierarchy through the 'active presencing' of these languages. Furthermore, Taylor and Kochem (2020) point out that technology has paved the way for new methods in language learning, but marginalised and/or indigenous populations often lack access to these tools. While this does not apply to all minoritised language groups, their critical literature review examines how researchers have tried to overcome this hurdle by reaching out to indigenous populations and provide access to emerging technologies. They found that indigenous populations often lack access to these tools, making it difficult for them to utilize them for language maintenance, growth, and preservation, recommending that additional initiatives – further inclusion of indigenous languages within, and increased funding for, mainstream linguistics research – are taken to further develop technology and marginalised populations' access to them.

However, it should be noted that the technologies being described here do not engage with the many different aspects of language itself. Language exists in contexts far beyond the spoken or written word, and that these too are minoritised languages. Sayers et al. (2021), for example, discuss minority and under-resourced languages, but also considers touch and signed languages such as braille and British Sign Language. They also discuss elements of sentiment, such as politeness or emotional expression. They argue that the inequality faced by minority languages is not just a question of gathering enough data. The data in question will be in exactly the same format as for bigger languages: audio recordings, automatically transcribed and perhaps tidied up by humans for AI to digest. But for sign languages, there is a much bigger hill to climb. Sign language is multimodal: it involves not only making shapes with the hands; it also relies heavily on facial expression, gesture, gaze, and a knowledge of the other signer's own life. All that represents a much greater technological challenge than teaching a machine to hear us and speak like us. For the Deaf community, for example, they argue that the emerging AI era is perhaps less promising. Similarly, Janutunen et al. (2021) discuss the prerequisites for the machine translation of sign languages. They take a linguistic approach to discuss a process that would be necessary for automated translation of sign languages and conclude that delivering a universal tool for sign language translation that is both robust and cost-effective is still not realistic. They suggest that while such translation tools may not be necessary for communities of people who rely on sign language to communicate, their omission from the development of AI-powered translation technologies risks widening communication gaps even further.

Responses

In light of the challenges facing the development of AI tools for minority languages, particularly in building sufficient corpora and addressing biases in existing technologies, some have argued that there is a pressing need for more collaborative efforts and the inclusion of minority language speakers in the design and production of AI technologies to ensure their effectiveness and cultural appropriateness. Yim (2024), for example, engages in ideas of indigenous data sovereignty to ask some key ethical questions about the capacity of communities to consent to their data being used when it is collected to train large language models. They argue that, while an early and undeveloped topic of discussion, the application of natural language processing and large language models for language representation raises a multitude of questions about future data consent. Just as prior linguistic data has been collected from speakers who were unaware of how it might be shared online, consent for continued usage of legacy data should not be assumed. Instead, it should be reconfirmed by the origin community for applications like machine learning. This is particularly pertinent, as research into the interaction between AI-powered translation and copyright law has highlighted the bias within AI translation systems and the impact that this has on marginalised communities (Yanisky-Ravid and Martens 2019). As such, the dominance of large firms in the AI sector results in the potential for the companies that develop translation systems to reflect the biases within their organisations, or within the communities that they aim to serve. This, in turn, can result in minority and marginalised communities having no access to systems of accountability when it comes to rectifying mistakes or addressing harms. Moreover, Rustagi (2021) argues that language and social reality are mutually reinforcing; as a result, natural language processing presents a unique opportunity to shift social reality at scale, advancing social justice by promoting linguistic justice. They provide an overview of how language and bias are intertwined and implications for building NLP tools that actively advance equity and inclusion. The authors suggest a positive framing with four layers of linguistic structure: The first three of these layers are concerned with the organisation of words and phrases into grammatical structure, and their use over time. After that, comes concerns about power inequities. They conclude that all NLP systems have the opportunity to support and advance linguistic justice by including and serving speakers of a wide variety of languages. Because language can provide access to power, ensuring that speakers of different language varieties are given equitable access is essential. Furthermore, NLP tools provide a unique possibility to achieve greater access to information cross-linguistically as these tools become increasingly accurate at tasks like translation and allow for a greater set of language varieties to circulate and gain the prestige and recognition that they deserve.

However, in order for linguistic justice to be promoted through the use of AI tools, there needs to be active efforts to ensure that such tools are being developed and implemented in environments that are able to support their effectiveness. In Wales, for example, a linguistic infrastructure policy has been put forward under the Cymraeg 50 framework that seeks to develop six areas of work: dictionaries; terminology; corpora; standardisation; centralised website; and a new unit to coordinate Welsh linguistic infrastructure. Two corpora are currently available: the National Corpus of Contemporary Welsh and the Welsh National Corpora Portal. Such a policy can serve to support the development of more Welsh language focused AI tools, but questions remain about access to this data and whether it is sufficient for user needs in the development of any new technologies.

More investment in infrastructure may be needed if minority languages are to be preserved in an AI era, particularly as adoption of such technologies become mainstreamed within public and work life. For minority language speakers, the growing use of ChatGPT in workplaces, for example, risks creating further barriers and exclusion, particularly in light of the speed at which these technologies are being developed and adopted.

References

AARP. (2019). Can Artificial Intelligence Outsmart Age Bias? *Online article*, available at: https://www.aarp.org/work/age-discrimination/ai-job-recruiting-age-bias/

Aisa, R., Cabeza, J., & Martin, J., (2023). Automation and aging: The impact on older workers in the workforce, *The Journal of the Economics of Ageing, Elsevier*, vol. 26(C).

Ajunwa, I. (2020). The "black box" at work. *Big Data & Society*, 7(2). https://doi.org/10.1177/2053951720938093

Ajunwa, I. (2023). *The Quantified Worker: Law and Technology in the Modern Workplace*. Cambridge, United Kingdom: Cambridge University Press.

Algorithm Watch. (2023). Cultural Hegemony: How Generative AI Systems Reinforce Existing Power Structures. https://sustain.algorithmwatch.org/en/cultural-hegemony-how-generative-ai-systems-reinforce-existing-power-structures/

Arisoy-Gedik, C., & Ceyhan, A. Ilkay. (2024). The Gender Gap in the Digital Era: Reaching Algorithmic Fairness and Technological Inclusivity in Network Society. *Communication Papers. Media Literacy and Gender Studies*., 13(26), 64–75. DOI: https://doi.org/10.33115/udg_bib/cp.v13i26.23038

Atkinson, J. and Collins, P. (2024). Algorithmic Management and a New Generation of Rights at Work. Institute of Employment Rights. Available at: https://ssrn.com/abstract=4853536

Bacon and Hoque, (2024). Disability at work: Briefing notes. Web page. Available at: https://www.disabilityatwork.co.uk/briefing-notes/

Bogen, M. & Rieke A. (2018). Help Wanted: An Exploration of Hiring Algorithms, Equity and Bias. *Technical Report. Upturn*. Available at: https://www.upturn.org/static/reports/2018/hiring-algorithms/files/Upturn%20--%20Help%20Wanted%20-%20An%20Exploration%20of%20Hiring%20Algorithms,%20Equity%20and%20Bias.pdf

Booth, R. (2021). Ex-Uber driver takes legal action over 'racist' face-recognition software. *Guardian*. October 5. Available at: https://www.theguardian.com/technology/2021/oct/05/ex-uber-driver-takes-legal-action-over-racist-face-recognition-software

Bowker, L. (2010). Can Machine Translation meet the needs of official language minority communities in Canada? A recipient evaluation. In: Vandepitte, S. (ed.) *Looking for Meaning: Methodological Issues in Translation Studies*. Antwerp: Artesis, pp. 123.

Brand, A. (2023). *Generative AI: New Wave of Workers Adopt Generative AI in Pursuit of Greater Job Satisfaction*. Online article HR Review, available at: https://hrreview.co.uk/hr-news/reward-news/generation-ai-new-wave-of-workers-adopt-generative-ai-in-pursuit-of-greater-job-satisfaction/373017

Brown, E. (2021). The Femtech Paradox: How Workplace Monitoring Threatens Women's Equity, 61 Jurimetrics pp. 289-329. Available at: THE FEMTECH PARADOX: HOW WORKPLACE MONITORING - ProQuest

Brussevich, M., Dabla-Norris, E., & Khalid, S. (2019), Is technology widening the gender gap? Automation and the future of female employment, *IMF Working Papers*, Vol. 19, No 91. Available at: Is Technology Widening the Gender Gap? Automation and the Future of Female Employment

Brynjolfsson, E., Li, D., & Raymond, L. (2023). Generative AI at Work, *National Bureau of Economic Research*. Available at: https://www.nber.org/papers/w31161

Burgess, J., & Connell, J. (2020). New technology and work: Exploring the challenges. *The Economic and Labour Relations Review,* 31(3), 310-323. https://doi.org/10.1177/1035304620944296

Buyl, M., Cociancig, C., Frattone, C. and Roekens, N. (2022). Tackling Algorithmic Disability Discrimination in the Hiring Process: An Ethical, Legal and Technical Analysis. ACM Conference on Fairness, Accountability, and Transparency (FAccT '22, New York, NY, USA, Association for Computing Machinery 21 June 2022), pp. 1071-1082. https://dl.acm.org/doi/pdf/10.1145/3531146.3533169?casa_token=G5Cac0vDA_IAAAAA:gnbV7nEeEOUY5KUm7A5Rppzv5ZB2BKpTRHd30U2NT6ab85VaU8DhKnl0ZzR51LL4vvrgtKqpRQ

Cant, C. (2020). The warehouse without walls: A workers' inquiry at Deliveroo, *ephemera: theory & politics inorganization*, 20(4). Available at: Microsoft Word - 20-4Cant LO.docx

Carlsson, M. & Eriksson, S. (2019). Age discrimination in hiring decisions: Evidence from a field experiment in the labour market, *Labour Economics, Elsevier*, 59(C):173-183.

Celentano, D. (2023). 'Be your own boss'? Normative concerns of algorithmic management in the gig economy: reclaiming agency at work through algorithmic counter-tactics, *Philosophy and Social Criticism*, 0(0) 1–24

Chen, J. Y. (2018). Thrown under the bus and outrunning it! The logic of Didi and taxi drivers' labour and activism in the on-demand economy, *New Media & Society* 20(8): 2691–2711. DOI: https://doi.org/10.1177/1461444817729149.

Claypool, C., Carey, A. C. and Hart, L. (2021). Centering Disability in Technology Policy: Issue Landscape and Potential Opportunities for Action. Report, pp.1-100. Centre for Democracy and Technology. Available at: https://cdt.org/wp-content/uploads/2021/12/centering-disability-120821-1326-final.pdf

Coleg Cymraeg Cenedlaethol. (2024). Technology and Bilingualism Conference. [Online] Available from: Technology and Bilingualism Conference | Coleg Cymraeg Cenedlaethol

Consultancy UK (2024). *Workers turning to AI to beef up their office skills*. Online article, available at: https://www.consultancy.uk/news/37444/workers-turning-to-ai-to-beef-up-their-office-skills

Corcoran, P., Palmer, G., Arman, L., Knight, D. and Spasic, I. (2021). Creating Welsh Language Word Embeddings. *Applied Sciences* 11(15) https://doi.org/10.3390/app11156896

Corvite, S., Roemmich, K. Rosenberg, T. I., and Andalibi, N. (2023). Data Subjects' Perspectives on Emotion Artificial Intelligence Use in the Workplace: A Relational Ethics Lens. Proceedings of ACM Human-Computer Interactions. 7, CSCW1, Article 124. 38 pages. https://doi.org/10.1145/3579600

Cunliffe, D. (2019). Minority Languages and Social Media. In: Hogan-Brun, G., O'Rourke, B. (eds) The Palgrave Handbook of Minority Languages and Communities. Palgrave Macmillan, London. https://doi.org/10.1057/978-1-137-54066-9_18

Cunliffe, D., Vlachidis, A., Williams, D. and Tudhope, D. (2022). Natural language processing for under-resourced languages: Developing a Welsh natural language toolkit. *Computer Speech and Language*. 72(1) https://doi.org/10.1016/j.csl.2021.101311

Dalingwater, L. (2018). Neo-liberalism and Gender Inequality in the Workplace in Britain, *Revue Française de Civilisation Britannique* DOI: https://doi.org/10.4000/rfcb.1802

Datta, A., Tschantz, M. C., & Datta, A. (2015) Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination, *Proceedings on Privacy Enhancing Technologies*, 92–112. Available at: https://www.andrew.cmu.edu/user/danupam/dtd-pets15.pdf

Deeptrace, (2019). The State of Deepfakes: Landscape, Threats, and Impact. Available online: https://regmedia.co.uk/2019/10/08/deepfake_report.pdf

Department for Business and Trade (2023). *Trade Union Membership, UK 1995-2022: Statistical Bulletin*, online article, available at: https://assets.publishing.service.gov.uk/media/646dcc477dd6e7000ca9b289/Trade_Union_Membership_UK_1995-2022_Statistical_Bulletin.pdf

Dubber, M. D., Pasquale, F., Das, S., Ajunwa, I., & Schlund, R. (2020). Algorithms and the Social Organization of Work [Bookitem]. In *The Oxford Handbook of Ethics of AI* (1st ed.). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190067397.013.52

EIGI (2021). 'Artificial intelligence, platform work and gender equality', European Institute for Gender Equality report, available at https://eige.europa.eu/publications-resources/publications/artificial-intelligence-platform-work-and-gender-equality?language_content_entity=en

Eisenstadt, L. (2022). #MeTooBots and the AI Workplace. *University of Pennsylvania Journal of Business Law*, Available at SSRN: https://ssrn.com/abstract=3921186

Engler, A. (2019). For some employment algorithms, disability discrimination by default. *Brookings*, October 31. Available at: https://www.brookings.edu/articles/for-some-employment-algorithms-disability-discrimination-by-default/

Equal Rights Trust. (2023). Discriminatory By Default? Case studies on the discriminatory impacts of algorithmic decision making systems. Pp.1-29. https://www.equalrightstrust.org/sites/default/files/ertdocs/Discriminatory%20by%20Default.pdf

European Network Against Racism. (2020). Artificial Intelligence in HR: How to address racial biases and algorithmic discrimination in HR? Report. Brussels. Pp. 1-28. https://www.enar-eu.org/artificial-intelligence-in-hr-how-to-address-racial-biases-and-algorithmic-discrimination-in-hr/

Farrar, J., Aslam, Y. and Carlo, S. (2021). Letter to Mayor Khan. March 18. Available at: https://www.workerinfoexchange.org/letter-to-mayor-khan

Ferdous, S. (2023). Are Older Workers Ready for an AI Takeover at Work? *The Oxford Institute of Population Aging*, Available at: https://www.ageing.ox.ac.uk/blog/Are-Older-Workers-Ready-for-an-AI-Takeover-at-Work

Fruchterman, J. and Mellea, J. (2018). Expanding Employment Success for People with Disabilities. Bentech report. Available at: https://benetech.org/wp-content/uploads/2018/11/Tech-and-Disability-Employment-Report-November-2018.pdf

Givens, A. (2020). How Algorithmic Bias Hurts People With Disabilities. *Slate*, February 6. Available at: https://slate.com/technology/2020/02/algorithmic-bias-people-with-disabilities.html

Ghio, A. (2024). Democratising academic research with Artificial Intelligence: The misleading case of language. *Critical Perspectives on Accounting* vol. 28.

Gomez-Herrera, E. and S. Koeszegi (2022) 'A gender perspective on artificial intelligence and jobs: the vicious cycle of digital inequality', Working Paper 15/2022, Bruegel.

Glazko, K., Mohammed, Y., Kosa, B., Potluri, V., and Mankoff, J. (2024). Identifying and Improving Disability Bias in GPT-Based Resume Screening. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (FAccT '24). Association for Computing Machinery, New York, NY, USA, 687–700. https://doi.org/10.1145/3630106.3658933

Gent, C. (2018). *The Politics of Algorithmic Management*. Thesis submitted for the degree of Doctor of Philosophy in Interdisciplinary Studies, University of Warwick.

Gerken, M. (2022). Facilitating the implementation of the European Charter for Regional or Minority Languages through artificial intelligence. Council of Europe Secretariat of the European Charter for Regional or Minority Languages.

Giakoumis, D., Votis, K. E Altsitsiadis, E., Segkouli, S., IPaliokas,I. & Tzovaras, D. (2019). Smart, personalized and adaptive ICT solutions for active, healthy and productive ageing with enhanced workability. *In Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '19). Association for Computing Machinery*, New York, NY, USA, 442–447. https://doi.org/10.1145/3316782.3322767

GMB (2021). *Historic workers' rights win: Supreme Court rules in Uber drivers' favour*, Online article by GMB union, available at: https://www.gmb.org.uk/news/uber-workers-rights-historic-gmb-supreme-court-drivers-legal-battle

Goggin, G., Ellis, K., & Hawkins, W. (2019). Disability at the centre of digital inclusion: assessing a new moment in technology and rights. *Communication Research and Practice*, 5(3), 290–303. https://doi.org/10.1080/22041451.2019.1641061

Gross, N. (2023) What ChatGPT Tells Us about Gender: A Cautionary Tale about Performativity and Gender Biases in AI, *Social Sciences* 12(8), DOI: https://doi.org/10.3390/socsci12080435

Hajric, E., Arevalo, F. N., Bruce, L., Smith, F. A. and Michael, K. (2024). Facial Emotion Recognition in the Future of Work: Social Implications and Policy Recommendations. *IEEE Transactions on Technology and Society*, doi: 10.1109/TTS.2024.3477512 https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10726575

Hekkala, S. & Hekkala, R.(2021). Integration of Artificial Intelligence into Recruiting Young Undergraduates: the Perceptions of 20–23-Year-Old Students. *Proceedings of the 54th Hawaii International Conference on System Sciences*. Available at: https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1006&context=hicss-54

Henry, M. et al. (2017) 'Learning a minority language through authentic conversation using an online social learning method', Computer Assisted Language Learning, 31(4), pp. 321–345. doi: 10.1080/09588221.2017.1395348. https://www.tandfonline.com/doi/epdf/10.1080/17439884.2023.2278111?needAccess=true

Hering, A. (2023). Indeed's AI at Work Report: The People Behind the Jobs GenAI is Most- and Least-Poised to Change — A Look at Age, Gender, and Race. Indeed Hiring Lab. October 26. https://www.hiringlab.org/2023/10/26/indeed-ai-at-work-report-people/

Houser, K. (2019). Can AI Solve the Diversity Problem in the Tech Industry? Mitigating Noise and Bias in Employment Decision-Making, SSRN. Available at: https://ssrn.com/abstract=3344751

IBA (2017). Artificial Intelligence and Robotics and Their Impact on the Workplace. *International Bar Association Global Employment Institute*. Available at: https://onwork.edu.au/bibitem/2017-Wisskirchen,G-Biacabe,B+T-etal-Artificial+intelligence+and+robotics+and+their+impact+on+the+workplace/

ILO. (2019). *Understanding the gender pay gap. In Women in Business and Management.*

IMF (2024). Gen-AI: Artificial Intelligence and the Future of Work. *International Monetary Fund.*

IPPR (2024). Transformed by AI: How generative artificial intelligence could affect work in the UK – and how to manage it. *Institute for Public Policy Research*. Available at: Transformed by AI: How generative artificial intelligence could affect work in the UK – and how to manage it | IPPR

IWGB. (2021). IWGB takes Uber to court over racist facial recognition as Black Lives Matter UK backs the drivers' strike and boycott on Wednesday 6 October. Blog post. October 5. Available at: https://iwgb.org.uk/en/post/racist-facial-recognition/

Janutunen, T., Rousi, R., Raino, P., Turunen, M. Valipoor, M.M., and Garcia, N. Is there any hope for developing automated translation technology for sign languages? (2021). In: Hamalainen, M., Partanen, N. and Alnajjar, K. (eds.) Multilingual Facilitation. Helsinki: Rootroo. Available at: IsThereHopeForAutomatedSLTranslation_final

Jones, D. (2022). Development and Evaluation of Speech Recognition for the Welsh Language. In Proceedings of the 4th Celtic Language Technology Workshop within LREC2022, pages 52–59, Marseille, France. European Language Resources Association. Available at: Development and Evaluation of Speech Recognition for the Welsh Language - ACL Anthology

Jung, C., & Desikan, B. (2024). 'Transformed by AI: How generative artificial intelligence could affect work in the UK – and how to manage it' Institute of Public Policy Research report, available at: https://ipprorg.files.svdcdn.com/production/Downloads/Transformed_by_AI_March24_2024-03-27-121003_kxis.pdf

Keegan, T. T., & Evas, J. (2012). Nudge! Normalizing the use of Minority Language Ict Interfaces. AlterNative: An International Journal of Indigenous Peoples, 8(1), 42-52. https://doi.org/10.1177/117718011200800104

Kellogg, K., Valentine, M., & Christin, A. (2020). Algorithms at Work: The New Contested Terrain of Control. *ANNALS*, 14, 366–410, https://doi.org/10.5465/annals.2018.0174

Kelly-Lyth, A. (2023). Algorithmic discrimination at work. *European Labour Law Journal*, 14(2), 152-171. https://doi.org/10.1177/20319525231167300

Kim, P. (2019). Big Data and Artificial Intelligence: New Challenges for Workplace Equality. *Scholarship@WashULaw*. 432. Available at: https://openscholarship.wustl.edu/law_scholarship/432

Kim, P. (2022). AI and Inequality in Kristin J., & Carla R., eds. *The Cambridge Handbook on Artificial Intelligence & The Law*, Cambridge University Press.

Kira, B. (2024). When non-consensual intimate deepfakes go viral: The insufficiency of the UK Online Safety Act, *Computer Law & Security Review: The International Journal of Technology Law and Practice*. DOI: https://doi.org/10.1016/j.clsr.2024.106024

Kohlrausch, B., &Weber, L. (2020). Gender Relations at the Digitalised Workplace: The Interrelation Between Digitalisation, Gender, and Work. *Gender a výzkum / Gender and Research* 21 (2): 13–31, DOI: http://dx.doi.org/10.13060/gav.2020.010.

Komp-Leukkunen, K, Poli, A, Hellevik, T, Herlofson, K, Heuer, A, Norum, R, Solem, PE, Khan, J, Rantanen, V & Motel-Klingebiel, A . (2022). 'Older Workers in Digitalizing Workplaces: A Systematic Literature Review', *The journal of aging and social change*, vol. 12, no. 2, pp. 37-59. Available at: Older Workers in Digitalizing Workplaces: A Systematic Literature Review | Semantic Scholar

KPMG (2024). *UK workers turn to social media and AI to sharpen office skills*. Online article. Available at: https://kpmg.com/uk/en/home/media/press-releases/2024/05/uk-workers-turn-to-social-media-and-ai-to-sharpen-office-skills.html#:~:text=Over%20half%20(56%25)%20of,how%20to%20use%20the%20technology

Kshetri, N (2024) Linguistic Challenges in Generative Artificial Intelligence: Implications for Low-Resource Languages in the Developing World, *Journal of Global Information Technology Management*, 27:2, 95-99, DOI: 10.1080/1097198X.2024.2341496

Lane, M. (2024). Who will be the workers most affected by AI?: A closer look at the impact of AI on women, low-skilled workers and other groups", *OECD Artificial Intelligence Papers*, No. 26, OECD Publishing, Paris, https://doi.org/10.1787/14dc6f89-en.

Lawrence, M. (2018) The forward march of robots halted? Automation, employment and inequality, *IPPR*. Available at: https://www.ippr.org/articles/the-forward-march-of-robots-halted-automation-employment-and-inequality

MacDonald, R., & Giazitzoglu, A. (2019) Youth, enterprise and precarity: or, what is, and what is wrong with, the 'gig economy'? *Journal of Sociology*, 55(4), 724-740. https://doi.org/10.1177/1440783319837604

Maliki, K., & Naji, F. (2024) Gender inequality in the sphere of artificial intelligence: Theoretical approach, *Journal of Autonomous Intelligence* 7(3), DOI: 10.32629/jai.v7i3.1394

Maurer, R. (2023). NYC Outlines Compliance Requirements for AI Bias Law. *Society for Human Resource Management*. April 16. Available at: https://www.shrm.org/topics-tools/news/talent-acquisition/nyc-outlines-compliance-requirements-ai-bias-law

Minority Rights Group International. (2020). Minority and Indigenous Trends 2020. Focus on Technology. Available at: https://www.ohchr.org/sites/default/files/Documents/Issues/Racism/SR/RaceBordersDigitalTechnologies/Minority_Rights_Group_International.pdf

Mohr, T. S. (2014). *Why Women Don't Apply for Jobs Unless They're 100% Qualified*. HBR

Muro, M., Maxim, R. & Whiton, J., (2019). *What Jobs are Affected by AI?* Report by Brookings Metropolitan Policy Program. Available at: What jobs are affected by AI? Better-paid, better-educated workers face the most exposure

Navani, R., Florisson, A. and Wilkes, M. (2023). The disability gap: Insecure work in the UK. Blog post. Available at : https://www.lancaster.ac.uk/work-foundation/publications/the-disability-gap-insecure-work-in-the-uk

Nugent, S., Jackson, P., Scott-Parker, S., Partridge, J., Raper, R., Shepherd, A., Bakalis, C., Mitra, A., Long, J., Maynard, K., and Crook, N. (2020). Recruitment AI has a Disability Problem: questions employers should be asking to ensure fairness in recruitment. Whitepaper, Institute for Ethical Artificial Intelligence, pp.1-18. Available at: https://nervetumours.org.uk/images/downloads/IEAI_Disability_Whitepaper_VIII_(1).pdf

No Language Left Behind. (2022). No Language Left Behind: Scaling Human-Centered Machine Translation. Meta AI, UC Berkely, Johns Hopkins University

PwC. (2024). PwC's 2024 *AI Jobs Barometer*. Report by PWC. Available at: AI Jobs Barometer | PwC

Pizzinelli, C. & Tavares, M. (2024). What the AI revolution means for jobs, productivity and global prosperity. *London School of Economics blog*. Available at: https://blogs.lse.ac.uk/inequalities/2024/05/29/what-the-ai-revolution-means-for-jobs/

Pizzinelli, P., Panton, A., Tavares, M., Cazzaniga, M., Li, L., (2023). Labour Market Exposure to AI: Cross-country Differences and Distributional Implications, *IMF Working Paper WP/23/216*, Available at: https://www.imf.org/en/Publications/WP/Issues/2023/10/04/Labor-Market-Exposure-to-AI-Cross-country-Differences-and-Distributional-Implications-539656

Pradhan, U. and Dey, J. (2023). Language, artificial education, and future-making in indigenous language education. *Learning, Media and Technology*.

Rhue, L. (2018). "Racial Influence on Automated Perceptions of Emotions." SSRN Electronic Journal, http://dx.doi.org/10.2139/ssrn.3281765

Rood, M. L., & Schriner, J. (2021). The Internet never forgets: Image-based sexual abuse and the workplace. In L. R. Salazar (Ed.), *Handbook of research on cyberbullying and online harassment in the workplace* (pp. 107–128). Business Science Reference/IGI Global.

Rosenblat, A. & Stark, L. (2016). Algorithmic Labour and Information Asymmetries: A Case Study of Uber's Drivers. *International Journal of Communication*, 10: 3758-3784.

Rustagi, I. (2021). Advancing social justice through linguistic justice: strategies for building equity fluent NLP technology. In Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '21), *October 5s9, 2021, s, NY, USA. ACM, New York, NY, USA, 9 pages.* https://doi.org/10.1145/3465416.3483301 Available at: (99+) Advancing social justice through linguistic justice: Strategies for building equity fluent NLP technology

Sánchez-Monedero, J., Dencik, L., & Edwards, L.(2020). What Does It Mean to 'Solve' the Problem of Discrimination in Hiring? *SSRN electronic journal*. Available at SSRN: https://ssrn.com/abstract=3463141

Sannon, S. and Cosley, D. (2022). Toward a More Inclusive Gig Economy: Risks and Opportunities for Workers with Disabilities. Proceedings of ACM Hum.-Comput. Interact. 6, CSCW2, Article 335, pp.1-31. https://doi.org/10.1145/3555755

Salas-Pilco, S.Z.; Xiao, K.; Oshima, J. (2022). Artificial Intelligence and New Technologies in Inclusive Education for Minority Students: A Systematic Review. Sustainability, 14, 13572. https://doi.org/10.3390/su142013572

Sayers, D., R. Sousa-Silva, S. Höhn et al. (2021). The Dawn of the Human-Machine Era: A forecast of new and emerging language technologies. Report for EU COST Action CA19102 'Language In The Human-Machine Era'. www.lithme.eu. Available at: (99+) The Dawn of the Human-Machine Era: A forecast of new and emerging language technologies

Schmidpeter, B., & Winter-Ebmer, R. (2018). How do automation and offshorability influence unemployment duration and subsequent job quality? IHS Economics Series, *No. 343, Institute for Advanced Studies (IHS), Vienna.*

Singh, S. & Narayanan R. (2023). Is ChatGPT Biased? A Review. DOI: https://doi.org/10.31219/osf.io/9xkbu

Somers, H. (2004). Machine translation and Welsh: the way forward / Cyfieithu Peirianyddol a'r gymraeg: y ffordd ymlaen. A report for the Welsh language board. Manchester: Centre for Computational Linguistics UMIST.

Stark, D., & Pais, I. (2021). Algorithmic Management in the Platform Economy. *Sociologica, 14*(3), 47–72. DOI: https://doi.org/10.6092/issn.1971-8853/12221

Tan, Y. and Jehom, W. (2024). The Function of Digital Technology in Minority Language Preservation: The Case of the Gyalrong Tibetan Language. *Preservation, Digital Technology & Culture*. https://doi.org/10.1515/pdtc-2024-0021

Taylor, J. and Kochem, T. (2020). 'Access and empowerment in digital language learning, maintenance, and revival: a critical literature review', Diaspora, Indigenous, and Minority Education, 16(4), pp. 234–245. doi: 10.1080/15595692.2020.1765769.

Tilmes, N. (2022). Disability, fairness, and algorithmic bias in AI recruitment. *Ethics and Information Technology*, 24:21. https://doi.org/10.1007/s10676-022-09633-2

Trosterud, T. (2006). Grammatically based language technology for minority languages. In: Saxena, A. and Borin, L. ed. Lesser-Known Languages of South Asia: Status and Policies, Case Studies and Applications of Information Technology. Berlin, New York: De Gruyter Mouton, pp. 293-316. https://doi.org/10.1515/9783110197785.3.293

TUC. (2016). Nearly two in three young women have experienced sexual harassment at work, TUC survey reveals, *Trades Union Congress*. Available at:. https://www.tuc.org.uk/news/nearly-two-three-young-women- have-experienced-sexual-harassment-work-tuc-survey-reveals

TUC (2020) Technology managing people - The worker experience, *Trades Union Congress report*, available at: https://www.tuc.org.uk/research-analysis/reports/technology-managing-people-worker-experience

Turner, J. (2024). NYC's Anti-Bias AI Hiring Law Has Largely Failed So Far. *Legal News and Analysis*. May 15. Available at: https://www.legalreader.com/nycs-anti-bias-ai-hiring-law-has-largely-failed-so-far/

UNESCO, OECD, &, IDB (2022) 'Report: The effects of AI on the working lives of women', *UNESCO, Paris*, https://doi.org/10.1787/14e9b92c-en

Unison (N.D.) Unison calling. A guide to organising in call centres. *Unison*, available at: https://www.unison.org.uk/content/uploads/2013/06/On-line-Catalogue208653.pdf

University of Edinburgh. (2023). AI initiative gives Gaelic a foothold in the digital age. [Online] Available at: https://www.ed.ac.uk/news/2023/ai-initiative-gives-gaelic-a-foothold-in-the-digit

Vasilescu M., Serban A,., Dimian G,, Aceleanu M., Picatoste X (2020) Digital divide, skills and perceptions on digitalisation in the European Union—Towards a smart labour market. *PLoS ONE* 15(4): e0232032. https://doi.org/10.1371/journal.pone.0232032

Welsh Government. (2018). Welsh Language Technology Action Plan. [Online] Available from: WG34015

Welsh Government. (2024). Empowering communities, strengthening the Welsh language. The Report of the Commission for Welsh-speaking Communities on communities with a higher density of Welsh speakers. [Online] Available from: 45614 Interim report by The Independent Commission on the Constitutional Future of Wales

Welsh Government. (2024). Generative AI – an overview. Research Briefing. Welsh Parliament: Cardiff. Available at: 24-14-generative-ai-en.pdf

West, M., R. Kraut, H. E. Chew. (2019) I'd Blush If I Could: Closing Gender Divides in Digital Skills through Education. *UNESCO*. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000367416

Whelan, A. (2021). Unequal Representation: Women in Clinical Research. Cornell Law Review Online. Available at SSRN: https://ssrn.com/abstract=4128449

Wilmers, N. (2024). Generative AI and the Future of Inequality. *An MIT Exploration of Generative AI, March*. https://doi.org/10.21428/e4baedd9.777b7123.

Wood, A. (2021). Algorithmic management consequences for work organisation and working conditions, *JRC Working Papers Series on Labour, Education and Technology, No. 2021/07, European Commission, Joint Research Centre (JRC), Seville.*

Wood, A. (2024). Algorithmic Management From Technology to Politics and Theory. Weizenbaum Journal Of The Digital Society. 4(3): pp.1-11. Available at: https://ojs.weizenbaum-institut.de/index.php/wjds/article/view/4_3_9/163

Worker Info Exchange (2021). Managed By Bots: Daat-Driven Exploitation in the Gig Economy. December. Pp. 1-82. Available at: https://5b88ae42-7f11-4060-85ff-4724bbfed648.usrfiles.com/ugd/5b88ae_8d720d54443543e2a928267d354acd90.pdf

World Economic Forum. (2024) Women are falling behind on generative AI in the workplace. *Here's how to change that*. Available at: https://www.weforum.org/agenda/2024/04/women-generative-ai-workplace/

Wright, H. (2024). Bates Wells and the EHRC supports Uber Eats driver who faced problematic AI checks. Bates Well blog post. March 26. https://bateswells.co.uk/updates/bates-wells-and-the-ehrc-supports-uber-eats-driver-who-faced-problematic-ai-checks/

Wright, L. Muenster, R, Vecchione, B., Qu, T., Cai, P. S., Smith, A., Comm 2450 Student Investigators, Metcalf, J., and Matias. J. N. (2024). Null Compliance: NYC Local Law 144 and the challenges of algorithm accountability. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24).* Association for Computing Machinery, New York, NY, USA, 1701–1713. https://doi.org/10.1145/3630106.3658998

Yanisky-Ravid, S. and Martens, C. (2019). From the Myth of Babel to Google Translate: Confronting Malicious Use of Artificial Intelligence – Copyright and Algorithmic Biases in Online Translation Systems. *Seattle University Law Review* 43(1) Available at: https://www.academia.edu/48730216/From_the_Myth_of_Babel_to_Google_Translate_Confronting_Malicious_Use_of_Artificial_Intelligence_Copyright_and_Algorithmic_Biases_in_Online_Translation_Systems

Yin, L., Alba, D. and Nicoletti, L. (2024). OpenAI's GPT Is a Recruiter's Dream Tool. Tests Show There's Racial Bias. *Bloomberg*, March 8.

Young, E., Wajcman, J., & Sprejer, L. (2021). *Where are the women? Mapping the gender job gap in AI*. Policy Briefing: Full Report. The Alan Turing Institute. Available at: https://www.turing.ac.uk/news/publications/report-where-are-women-mapping-gender-job-gap-ai

Zaniboni, S., Kmicinska, M., Truxillo, D. M., Kahn, K., Paladino, M. P., & Fraccaroli, F. (2019). Will you still hire me when I am over 50? The effects of implicit and explicit age stereotyping on resume evaluations. *European Journal of Work and Organizational Psychology*, 1-15. https://doi.org/10.1080/1359432X.2019.1600506

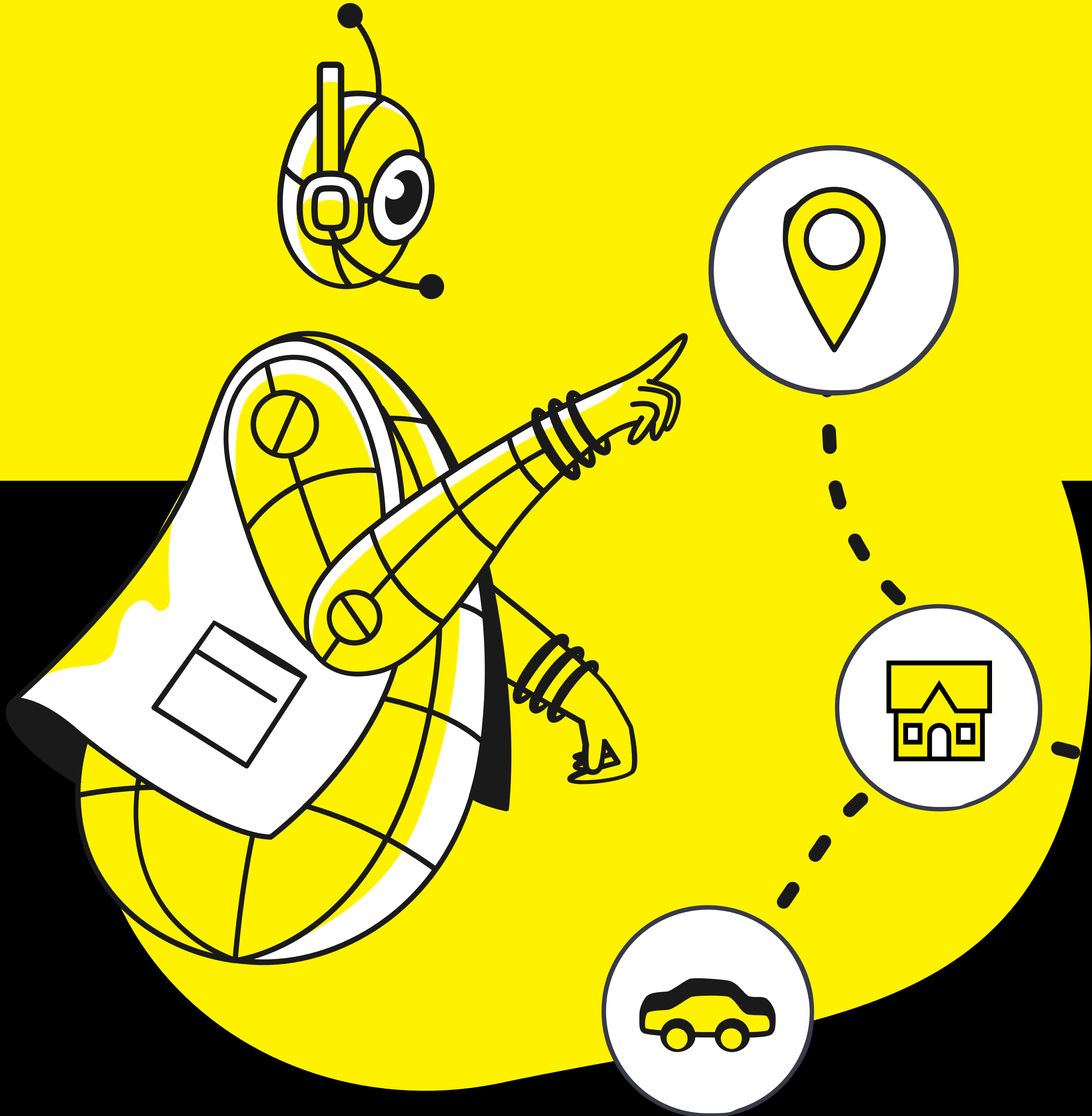
Zapata, D. (2021). New study finds AI-enabled anti-Black bias in recruiting. Thomson Reuters. June 18. https://www.thomsonreuters.com/en-us/posts/legal/ai-enabled-anti-black-bias/

Zhang, C., Tao, M., Huang, Q., Lin, J., Chen, Z., Feng, Y. (2024). MC2 : Towards Transparent and Culturally-Aware NLP for Minority Languages in China. Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 8832–8850 August 11-16, 2024 ©2024 Association for Computational Linguistics. Available at: https://aclanthology.org/2024.acl-long.479/

Zhuang, K. V. and Goggin, G. (2024). New possibilities or problems for disability and inclusion? The case of AI and ADMs across work. *Telematics and Informatics*, Volume 92, pp.1-10. ISSN 0736-5853, https://doi.org/10.1016/j.tele.2024.102156.

Zur, O. & Zur, A. (2011). On Digital Immigrants and Digital Natives: How the Digital Divide Affects Families, Educational Institutions, and the Workplace. *Zur Institute - Online Publication*. Available at: https://www.researchgate.net/publication/267213208_On_Digital_Immigrants_and_Digital_Natives_How_the_Digital_Divide_Affects_Families_Educational_Institutions_and_the_Workplace

Zyskowski, K., Morris, M. R., Bigham, J. P., Gray, M. L., & Kane, S. K. (2015). Accessible crowdwork? Understanding the value in the challenge of microtask employment for people with disabilities. *CSCW '15: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, pp. 1682 - 1693*. https://doi.org/10.1145/2675133.2675158



Data
Justice
Lab

TUC
Cymru