The Artificial Recruiter: Risks of Discrimination in Employers' Use of AI and Automated Decision-Making

Larsson, Stefan; White, James Merricks; Ingram Bogusz, Claire

Social Inclusion
2024 • Volume 12 • Article 7471
https://doi.org/10.17645/si.v12.7471

ARTICLE

Open Access Journal

The Artificial Recruiter: Risks of Discrimination in Employers' Use of AI and Automated Decision-Making

Stefan Larsson [1], **James Merricks White** [1], **and Claire Ingram Bogusz** [2]

[1] Department for Technology and Society, Lund University, Sweden
[2] Department of Informatics and Media, Uppsala University, Sweden

Correspondence: Stefan Larsson (stefan.larsson@lth.lu.se)

Submitted: 31 July 2023 **Accepted:** 26 February 2024 **Published:** 18 April 2024

Issue: This article is part of the issue "Artificial Intelligence and Ethnic, Religious, and Gender-Based Discrimination" edited by Derya Ozkul (University of Warwick), fully open access at https://doi.org/10.17645/si.i236

Abstract

Extant literature points to how the risk of discrimination is intrinsic to AI systems owing to the dependence on training data and the difficulty of post hoc algorithmic auditing. Transparency and auditability limitations are problematic both for companies' prevention efforts and for government oversight, both in terms of how artificial intelligence (AI) systems function and how large-scale digital platforms support recruitment processes. This article explores the risks and users' understandings of discrimination when using AI and automated decision-making (ADM) in worker recruitment. We rely on data in the form of 110 completed questionnaires with representatives from 10 of the 50 largest recruitment agencies in Sweden and representatives from 100 Swedish companies with more than 100 employees ("major employers"). In this study, we made use of an open definition of AI to accommodate differences in knowledge and opinion around how AI and ADM are understood by the respondents. The study shows a significant difference between direct and indirect AI and ADM use, which has implications for recruiters' awareness of the potential for bias or discrimination in recruitment. All of those surveyed made use of large digital platforms like Facebook and LinkedIn for their recruitment, leading to concerns around transparency and accountability—not least because most respondents did not explicitly consider this to be AI or ADM use. We discuss the implications of direct and indirect use in recruitment in Sweden, primarily in terms of transparency and the allocation of accountability for bias and discrimination during recruitment processes.

Keywords

ADM and risks of discrimination; AI and accountability; AI and risks of discrimination; AI and transparency; artificial intelligence; automated decision-making; discrimination in recruitment; indirect AI use; platforms and discrimination

1. Introduction

Perhaps the most talked about example of discrimination in AI-supported recruitment is Amazon's now-defunct résumé scanning tool that employed automated decision-making (ADM). Developed between 2014 and 2018, Amazon trained the tool on the credentials of previously recruited candidates to effectively identify and rank qualified applicants. However, the system—which largely looked for patterns—began downgrading résumés from female candidates. Although applicants' gender was not explicitly given, the AI system used indirect markers, such as "captain of the women's chess club," as proxies. After discovering this bias, Amazon's engineers tried to fix the problem by directing the system to treat these terms in a "neutral" way. The company eventually abandoned its efforts: They could not prevent the tool from discriminating based on gender and faced ongoing public criticism (cf. Myers West et al., 2019). This highlights one AI development challenge: the reliance on something already established on which to train (Larsson et al., 2023b).

In the above example, prevailing inequality in the tech industry has been argued to explain why Amazon had such a hard time developing a recruitment tool that did not reproduce that bias (Ajunwa, 2023). The problem thus lies partly in how markers for gender, ethnicity, and age are implicit in the recruitment environments, and partly that AI systems, due to their training data, risk reflecting historical or present biases instead of employers' stated requirements for professional skills. It is certainly not new knowledge that language can be gendered, i.e., that different words can have culturally established male or female associations that, for example, affect how job advertisements are perceived (cf. Gaucher et al., 2011). Through machine learning, though, these associations are often captured and hidden in an AI model (Leavy, 2018), mirroring stereotypes (cf. Larsson et al., 2023b). While well-documented in gendered language, proxies for race have also been seen to influence search engine algorithms (Noble, 2018).

It is not possible to understand a general trend through examination of one exemplar case, like the Amazon résumé scanning tool. Research shows that AI and ADM are increasingly used in worker recruitment, specifically through LinkedIn (Ajunwa & Greene, 2019; Laukkarinen, 2023; Ruparel et al., 2020). However, this does not reveal how recruiters perceive their own reliance on such tools or their underlying data. To better understand the degree to which recruiters are aware when they are using AI, and the implications of this use, we turned to collecting data about how AI and ADM are used in Sweden. Considerable public attention and academic research have been paid to the use of algorithms in the gig economy (e.g., Jones, 2021; Woodcock & Graham, 2020) including in the Nordic region (Ilsøe & Söderqvist, 2022; Newlands, 2022), as well as broader fears of mass unemployment due to automation and robotisation (e.g., Pasquale, 2020). But these are not the only ways that AI and ADM are affecting working life. As an extension of widespread and ongoing social processes of digitisation and datafication (see Thylstrup, 2018; van Dijck, 2014), these tools have also found their way into regular practices of hiring and firing, and organising and monitoring workers (Lomborg, 2022) in a near-infrastructural way (Plantin et al., 2016). This platformisation (van Dijck et al., 2018) of anything automated in recruitment leads to a question of whether AI and ADM are directly used in-house by the professionals, or if it is indirectly used via digital services and platforms like LinkedIn (cf. Komljenovic, 2019). Recent research has attended to expectations, perceptions, and outcomes of hiring algorithms (Dencik & Stevens, 2021; Zhang & Yencha, 2022) and their inherent biases and developer attempts to mitigate them (Kelly-Lyth, 2021; Sánchez-Monedero et al., 2020). The mentioned distinction between direct and indirect uses of AI and ADM is relevant in terms of awareness of risks of discrimination needed for assuring accountability for preventive measures.

The heterogeneous deployment of AI and ADM in the workplace is a significant challenge for regulators. To the extent this deployment is dependent on large external platforms' services, the recently implemented Digital Services Act (DSA; see European Commission, 2022) is of interest, as it imposes transparency obligations on the largest tech firms and their services (including LinkedIn and Facebook). Furthermore, given the classification of employment and worker management as high-risk areas in the EU's forthcoming AI legislation (Veale & Zuiderveen Borgesius, 2021; on the provisional trilogue agreement see also European Commission, 2023), the risk for bias and discrimination in this area is of particular importance in highly digitised EU economies, such as Sweden. Lastly, the purpose of the Swedish Discrimination Act (SFS 2008:567) is to counter discrimination and in other ways promote equal rights and opportunities regardless of the basis of discrimination. Within working life and education, the law also requires employers and education providers to work continuously to promote equal rights and opportunities. There is, however, a challenge of how to interpret established legislation on anti-discrimination in the light of AI and ADM practices (cf. Wachter et al., 2021).

Given the risk of discrimination that seems inherent in the use of these tools, methods, and services, this study's research objective is to explore to what extent and in what contexts employers use AI and ADM for worker recruitment. It particularly engages with questions related to individual and organisational risk awareness, and what this implicates in terms of transparency and accountability for those risks. In what follows, we give some theoretical background to our study (Section 2), before describing our data and methods (Section 3). We then present and analyse the results of our questionnaires with representatives from 10 major recruitment agencies and 100 private sector companies with more than 100 employees (Section 4). We discuss these results in light of the research objective, as well as its implications for AI, ADM, and risk and discrimination (Section 5).

2. Theoretical Background

2.1. Definitional Difficulties

Coined in the 1950s, "artificial intelligence" has proven an elusive concept, both as a field of research (Hagendorff & Wezel, 2020) and in regulatory efforts by European legislators (European Commission, 2021). AI has, for instance, been defined narrowly to include only algorithms that exhibit human-like intelligence—including only algorithms that are self-referential and thus teach themselves (Berente et al., 2021)—and broadly as any kind of automated algorithm, including rule-based systems like decision trees (as per the European Commission's proposal for the AI Act in 2021). According to legal experts Gasser and Almeida (2017, p. 59), "AI is not a single technology, but rather a set of technologies and sub-disciplines ranging from areas such as speech recognition and computer vision to attention and memory, to name just a few." Although the field has developed a great deal in recent years, including paying attention to large language models and what has come to be referred to as "generative AI" or "general purpose AI," there is still much heterogeneity. This is partly a definitional problem, because processes can be automated and use AI in different ways, and a problem with the systems' lack of transparency. This provokes both an object of scrutiny and a methodological question: How are AI-associated risks perceived and mitigated when this concept is blurry and how can we study these risks through questionnaires without defining the concept in detail, respectively? In the following sections, we describe our broad perspective on AI and ADM and use it to explain our approach to these challenges.

Both to avoid getting bogged down in definitional issues and to explore AI in its broadest sense, we treated the definition of AI (and its implicit ambiguities) as part of the research inquiry. Terminological ambiguity around the concept of AI has implications for the questions addressed by this study. We see two challenges here: one methodological and one analytical. The methodological challenge is that we cannot expect our respondents to have a uniform understanding of what constitutes or does not constitute AI. The analytical challenge stems from the wide range of technologies and methods that fall under the umbrella terms "artificial intelligence" and "automated decision-making," and the difficulty of knowing when and where an AI solution is used within a broader system. These issues were handled by offering inclusive definitions of AI and ADM to participants and asking questions relating to both their use and awareness of their use:

> On behalf of the Equality Ombudsman, the use of artificial intelligence and other automated decision-making in recruitment processes in working life is investigated here. Automatic decision-making encompasses algorithmic and automated decision-making processes both with and without artificial intelligence. In the term, we include both fully automated decision-making and automated processes used as decision support. (translated from Swedish)

Similar definitional issues can be identified concerning ADM. In studies of public administration, the term "automated decision-making" is often used as a holistic concept to refer to various forms of automation in decision support (Roehl, 2022). Simpler rule-driven algorithmic systems for decision-making can thus be included; this means there is a longer history of the use of automated decisions, which may have been made daily but can still lead to challenges, as in the case of municipal income support (Kaun, 2022). However, the difference between automation for *decision support*, where a human has the final say, and fully automated *decision-making* (with no human intervention) can be crucial, not least because studies on the type of automation in the public sector that people generally trust show that support is far greater for ADM as a non-fully autonomous decision-support mechanism, where a human makes the final decision, than as a fully automated practice (Insight Intelligence, 2022). This type of literature often points to the significance and consequences of the development towards the quantification, or so-called datafication, of a range of processes, such as in welfare (Kaun et al., 2023) or in quantifying work and workers (Ajunwa, 2023).

2.2. The Risks of Discrimination of AI and ADM

AI and ADM are sometimes promoted as an unbiased alternative to human judgement (analysed in terms of "regimes of justification" by Dencik & Stevens, 2021). The argument goes that because machines do not harbour explicit or subconscious preferences for individual traits or characteristics, or because they are not prejudiced against groups of people, they cannot unfairly discriminate between them. Related to this is the assertion that because machines do not grow tired or dissatisfied with their work, they are unable to make the same mistakes or errors in judgement that humans often make.

There is now a significant body of work that shows the limits of this argument (e.g., Benjamin, 2019; Crawford, 2021; Noble, 2018). Because AI and ADM are developed by humans and trained on data that reflect human language and society, they internalise human biases and inequalities and come to reflect pre-existing problems in prevailing social structures (D'Ignazio & Klein, 2023; Larsson, 2019; Larsson et al., 2023b). Complex AI systems, like the recommender systems in search engines, can encode biases that are deeply embedded in historic and persistent distributions of opportunity. When internet studies scholar Noble (2018) searched for "black girls," she was astonished to discover that the first result returned by Google was for a pornography site. Her subsequent research into search results for websites, images, and geographical locations revealed systematic lack of credible and reliable information about women and people of colour, reinforcing racist and oppressive stereotypes, and the structural relations of domination that underlie them. Only when her work became a controversy for Google did they do anything to amend their algorithms and search results.

But it is not only complexity that is at issue here. Large language models trained on massive volumes of unstructured data enclose statistical relationships between words and groups of words within neural networks that have been argued to be inherently opaque (Peters, 2023). Just as the training data for these systems is rife with human prejudice, so do the systems themselves become prejudiced. When researchers asked such a system to complete the sentence "Man is to computer programmer as woman is to ________" its response was "homemaker" (see Bolukbasi et al., 2016; Caliskan et al., 2017). These kinds of problematic associations are the very same issues that beset OpenAI's ChatGPT, Google's Bard, and Microsoft's Bing (cf. Ferrara, 2023). They cannot easily be removed from trained models, and so developers have taken to placing evolving controls on model outputs. Even if prejudiced statements are fully and effectively removed from the training data, the output of large language models will remain inconsistent and unpredictable, and so always at risk of transgressing widely held values (Bender et al., 2021).

These kinds of problems also affect computer vision systems. The AI models at the core of these systems are trained on datasets of labelled images or videos. These data have often been produced by thousands of individuals who bid on small, on-demand jobs posted to crowdsource work marketplaces such as Amazon's Mechanical Turk—part of the gig economy touched on in the introduction section. While this allows any single data point to be cross-checked by several persons, audits of datasets that have been produced in this manner still reveal problems and inconsistencies (Crawford, 2021). These occur due to commonly repeated errors and biases in the tagging, and due to the unequal distribution of images within the dataset. In their study of facial recognition systems, Buolamwini and Gebru (2018) found that the lack of representation of female and dark-skinned faces within training data meant that the resultant models performed more poorly on women and people of colour. These results were compounded by intersecting absences and Black women were the most poorly recognised of all.

It is, however, far from clear that such problems with AI-driven image detection can be solved by simply fixing the data. In 2015, Google's photo web app automatically added "gorilla" tags to the faces of an African American couple. This was an unintended and embarrassing error that Google was quick to respond to by removing the tag (Russell, 2019). The problem is not only that the AI performed poorly on black faces or that the choice of the tag was racist to begin with; the problem is also that the meaning of this error can only be understood in terms of the history of American slavery and racial oppression, the justification for which was through scientific racism and the dehumanisation of Black people (Benjamin, 2019). This is not something that the model can know. The context of the error, whatever its cause, escapes the machine entirely.

The causes and likelihood of algorithmic bias and discrimination within complex technical systems are often hard to pin down. This is because such systems are not only complex, they also include trade secrets, and interests which do not wish to be (or should not be) openly auditable. Or, as Morondo Taramundi (2022, p. 73), articulates it: "Technical complexity, together with proprietary interest and economic calculations, makes it very difficult to understand exactly how algorithms discriminate." This is a particularly relevant concern for

AI and ADM used in hiring related to the large-scale tech platforms that many depend on in their indirect use of these technologies. Issues related to transparency and accountability are recognised challenges for AI systems and ADM as such (Larsson & Heintz, 2020), which can pose problems both for the governance of AI systems (e.g., Novelli et al., 2023; Taeihagh, 2021), as well as for large digital platforms (Geng, 2023; Kim & Moon, 2021). We return to this in Section 5. This research examines this accountability challenge linked to proprietary platforms and the associated knowledge gap when it comes to the use of AI and ADM as tools, methods, and services.

3. Methods and Data

The research was formulated by two of the authors in collaboration with the research company Novus and the Swedish Equality Ombudsman. In preparation for the questionnaire phase of the study, six unstructured scoping interviews were conducted with area experts. These helped clarify the questions to be included in the telephone questionnaire aimed at the two target groups: recruitment agencies and companies with more than 100 employees. The questionnaires were conducted via telephone calls to access an occupied group of professionals who are unlikely to take the time to answer text-based questionnaires sent to them via email (or mail). While the questionnaire responses entail self-reported data, we have no reason to believe that the respondents were biased or misleading in their answers.

The number of contact attempts per person for both groups was eight to ten. This strategy was motivated by the fact that this type of respondent is relatively unavailable, the population relatively small, and it was extremely important to get hold of the right respondent. The interview time for recruitment companies was on average 10 minutes and 12 minutes on average for the major employers.

3.1. Target Group 1: Recruitment Agencies

For recruitment agencies, we approached 50 of Sweden's largest recruitment agencies. Many recruitment agencies seem to have a policy of not responding to requests for information on how the services they offer are carried out. Of the 50 recruitment agencies contacted, 10 agreed to answer the questionnaire. Of these 10, nine recruit more than 50 people per year, and one recruits 11–20 people per year.

To ensure respondents had sufficient knowledge to answer the questions, the interviews were primarily conducted with chief technology officers or an equivalent HR manager. We started from a list provided by employer organisation Kompetensföretagen (part of Almega) of the 50 largest employment agencies (which includes recruitment agencies) in Sweden based on turnover. This list is published every quarter and we used the most recent edition.

3.2. Target Group 2: Major Employers

For major employers, 100 major Swedish employers responded to our questionnaire. This sample was comprised of 50 representatives of companies with 100–249 employees and 50 representatives of employers with more than 250 employees, of which 23 have more than 500 employees (see Table 1).

Table 1. Distribution of sample of major employers.

Number of employees in each participating company	Number of questionnaires answered
100–149 employees	22
150–249 employees	28
250–499 employees	27
500+ employees	23
Total	100

Statistics Sweden estimates there to be 1,904 companies in the range of 100–249 employees and 1,205 in the range of 250 employees and above (Statistics Sweden, 2023). Of the companies within the sample, about half (49%) recruit more than 50 new employees per year. Given the population size, a sample of 100 respondents gives more than 95% confidence in a 10% margin of error in the responses. It was determined that this would be sufficient for the purpose of this research, which was not to establish strong statistical relationships but to map potential discrimination risks. The data did not include any personal identifiers or personal information, only reported use by a particular organisation. This study was designed in collaboration with the public authority tasked with battling unlawful discrimination, the Swedish Equality Ombudsman, in a way that it should adhere to principles of ethical conduct in research as they are regulated in Sweden: In Sweden, existing law on ethical conduct in research, specifically the Ethics Review Act (but often in combination with the GDPR), aims to protect individual research subjects and respect for human dignity. The law defines what kinds of research must be approved by the Ethics Review Authority (Etikprövningsmyndigheten) before being carried out (Articles 3 and 4). The present study clearly does not meet this threshold and therefore did not require approval from the Ethics Review Authority. Furthermore, the Ethics Review Authority has recently issued guidance on what risks to consider before conducting research (Görman & Etikprövningsmyndigheten, 2023). Of particular relevance for this study is the information given to the respondents and the necessity to protect the integrity of the respondent through the anonymisation of personal data. In keeping with the Ethics Review Act and this supplementary guidance, independent ethical approval was not needed to conduct this study.

The questionnaire was geared towards HR managers or other senior managers with clear recruitment responsibilities or equivalent positions. Researchers took care to ensure that they were speaking to the person responsible and were often transferred several times within a company until they were put in contact with someone suitable. While these representatives may not accurately depict overall sentiment within the company, they are expected to have a greater level of awareness of recruitment practices and technologies than the company average. As these respondents act as a proxy for their company, care is taken to limit the analysis and not overstate the conclusions.

3.3. The Questionnaires

The two respondent groups received similar questionnaires. For the recruitment agencies, there were eight questions, with five follow-up questions depending on what the respondents answered. For the major employers, there were nine questions with four follow-up questions. The most important questions for this article can be found in Figures 1 through 4.

4. Results and Analysis

The analysis combines the descriptive results from the surveys with the theoretical needs required to analyse the importance of AI and automation concepts and how risks of discrimination can interact with technical solutions and in relation to already established knowledge in the field. The analysis thus includes the differences between perceived and reported use, direct and indirect use, and the risks associated with a lack of transparency in both AI systems and the services of large-scale platforms, along with questions about the allocation of accountability for any resulting bias or discrimination.

4.1. Do the Respondents Really Not Use AI?

Recruitment agencies were asked about the presence of these technologies in services offered to clients. Levels of perceived use of ADM within recruitment management systems were low, with eight out of ten responding that usage was to a fairly low extent, a very low extent, or not at all. Similar trends applied to the use of AI, where eight out of ten respondents said that use was very low or not at all. Among the major employers, 87% responded that the use of ADM/decision support in recruitment was fairly low or very low (Figure 1).

Given that respondents are expected to have slightly different perceptions of what constitutes AI and ADM, we cannot assume that this is an accurate representation of the actual use of these technologies. Rather, the results offer an overview of their *perceived* use amongst company representatives, who are assumed to have higher-than-average awareness levels. When we asked about specific applications and services, reported use was much higher.

4.2. Direct and Indirect Use of AI and ADM

Further questions in both the recruitment agency questionnaire (Figure 2) and in the major employer questionnaire were designed to better measure the actual use of AI and ADM (see Figure 3), and ultimately also revealed a distinction between direct and indirect use of AI and ADM. When recruiters were asked if

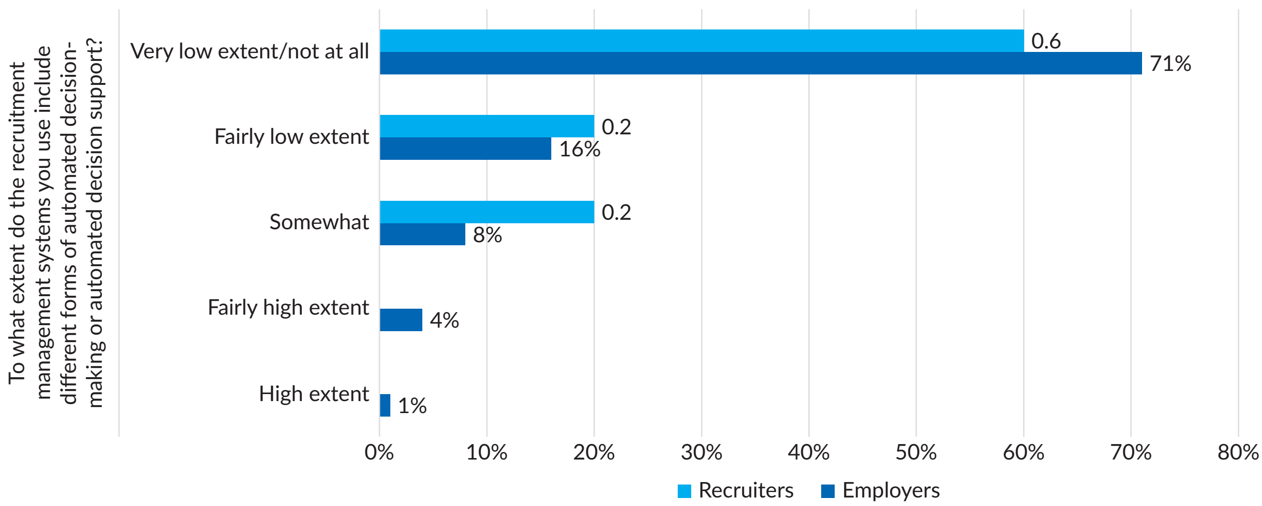

Figure 1. Do recruitment agencies and major employers use AI in their recruitment?

they used AI or another ADM to automatically match job seekers with jobs, seven out of ten of the respondents confirmed that they did (Figure 2). Since this is a clear reference to the variants of social media that, according to several studies, are common in job-matching contexts, such as LinkedIn, we can conclude that the respondents use AI functionalities with a large element of automation in their recruitment process, although they do not see it as such—hence the lower perceived use. Eight out of ten also said that they received information about suitable candidates from social media profiles, such as Facebook, LinkedIn, Instagram, TikTok, Twitter, and others (see Figure 2), which also make use of AI in their search and recommender operations. Most recruitment agencies (9 out of 10) tested candidates through games, IQ tests, or personality tests.

All, however, indicated that they used existing job platforms, such as the Swedish Public Employment Service or Monster. These job platforms offer search functions that can rank hits against either job postings (for job applicants) or jobseekers' résumés (for recruiters), using varying degrees of automation. So, although the agencies do not explicitly report this in the question about perceived AI use, all of them do, in fact, use some kind of ADM in their recruitment work, often through a third-party service.

Similarly, respondents in the large company questionnaire reported extensive use of digital platforms in their hiring practices, including recruitment platforms (80%) such as ReachMee, Varbi, LinkedIn Recruiter, and Teamtailor, and social media (55%), e.g., Facebook, LinkedIn, Instagram, and similar (see Figure 3).

Outsourcing of recruitment is also very common. 69% of the Swedish employers stated that they outsource some of their recruitment processes to a recruitment agency (see Figure 3). While the vast majority report that they use a recruitment platform and about half that they use social media in their recruitment processes, it is possible that some respondents interpreted outsourcing as one of these services. The extent to which and in what combination the respondents use these different services is not clear from the data.

The analytical challenge, which is related to the ambiguity involved in defining both AI and ADM, is visible in these questionnaire results. While it appears as though ADM is frequently used among recruiters (even if this is not fully acknowledged), it is unclear whether this is true of major employers. Recruitment management

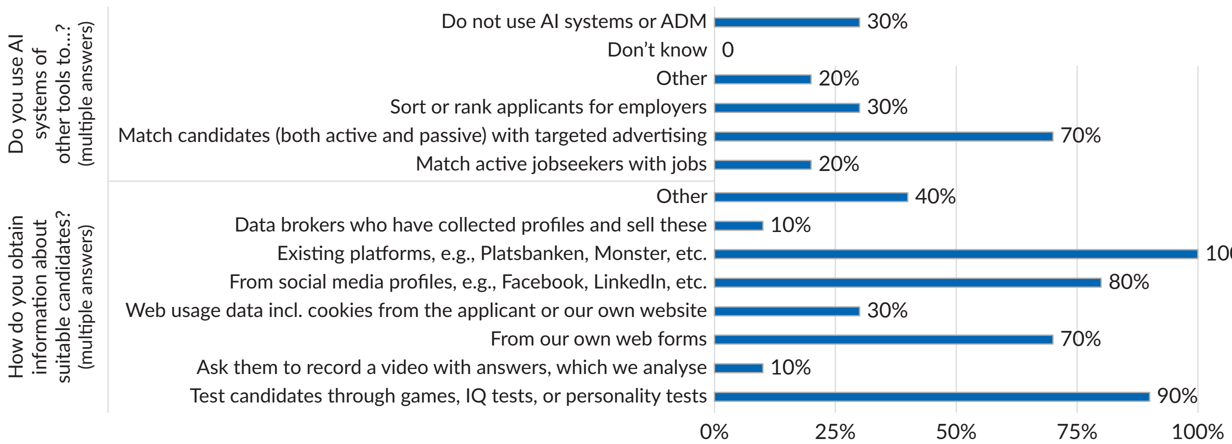

Figure 2. What do recruitment agencies use AI systems for, and how do they find information about candidates?

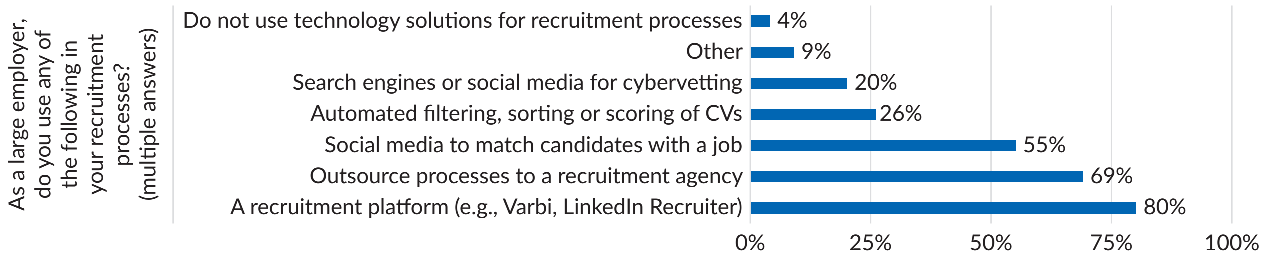

Figure 3. Which tools do major employers use in their recruitment processes?

platforms, such as ReachMee and Varbi, are integrated solutions developed to streamline job posting and candidate management. They are not marketed as AI-powered solutions but use forms of automated analysis for decision support. They can include forms of ADM, as we have defined the term. LinkedIn Recruiter, for instance, has publicised its use of AI to match candidates to jobs and to help identify potential recruitment opportunities (LinkedIn, 2018).

4.3. Accountability

Here, we use transparency to refer to auditability, access to databases, end-users' information skills, or the need for documentation about the algorithms used, and the goals governing the AI model (e.g., Larsson & Heintz, 2020; Larsson et al., 2023a). When the recruitment agencies were asked whether they told candidates that they used AI or ADM tools, six declined to answer, one said that they did not declare this but probably should, while three said that they did not know. This is clearly a challenge—compounded by both the ubiquity and extensive indirect use of tools with varying degrees of automation.

When the major employers were asked who they believe is responsible for discrimination occurring because of the use of an AI system or ADM, 84% stated that the responsibility lies with the company using the system (see Figure 4). At the same time, there is a clear and immediate risk that neither companies nor recruiters have any means of checking how the AI systems that are indirectly used ensure that they do not discriminate when matching candidates via social media or third-party recruitment services. The issue is complicated both by the inherent challenges in auditing AI systems and by the low transparency of platform companies (although the latter may change under the DSA, see Söderlund et al., 2024).

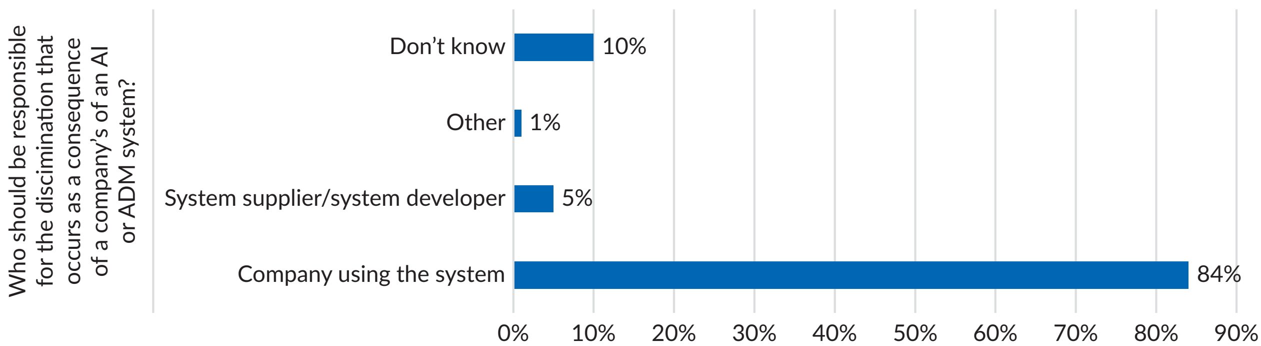

Figure 4. Who is accountable?

Overall, this data points to the extensive use of AI and ADM tools for recruitment both among recruitment agencies and among major employers in Sweden. It is interesting to note that recruiters (8 out of 10) and major employers (77%) say that their use of AI and ADM is low or very low, but that they make extensive use of tools and platforms like LinkedIn Recruiter and similar that explicitly market themselves as AI-enabled. Transparency and responsibility remain largely hypothetical concerns: Although major employers believe that they are themselves responsible for any discrimination, both extensive indirect use and laws like the DSA and the forthcoming AI Act suggest that the responsibility of platforms providing automated tools is not to be ignored.

5. Discussion

Taken together, these results indicate that there is low awareness of the everyday use of AI and ADM in the Swedish workplace by recruiters. This may indicate a lack of attention to the role played by these technologies in tools used every day. Every time a web search is performed (e.g., through Google), an automated recommendation system is used. Every time a user scrolls through someone's social media feed (e.g., on Facebook), an AI system determines what they see. Every ad that is seen on these websites and digital platforms has been selected by a complex automated model.

In this study, the difference between perceived and reported use of AI is striking and perhaps a testament to the ubiquity of AI tools in the workplace today. This study supports prior work that suggests that the increased prevalence of AI and other ADM in the workplace is not likely to lead to increased awareness. Instead, ubiquity leads to technologies being taken for granted and receding into the background (cf. Kaun, 2022; Lomborg, 2022). This growing body of evidence suggests there is a risk that these technologies will integrate into everyday life and not only become unreflexively trusted by the people who use them, but even ignored.

5.1. Indirect Use of AI and ADM

A further case in point lies in the distinction we made above between direct and indirect use of AI and ADM. Here, we understand direct use to occur when solutions have been significantly developed by the companies that use them, and indirect use to occur when a technical solution like AI is packaged into a service used by a company but provided by a third party, e.g., LinkedIn.

Indirect use is most apparent when one considers the use of technology solutions in recruitment processes. For instance, the finding that 26% of major employers said they used automated résumé filtering, sorting, or scoring (see Figure 3). Although this option was intended to measure direct use, it is more likely to be a mixture of both direct and indirect use. Much more common is that companies use a recruitment platform (with 80% of respondents saying that they do; see Figure 3), into which this type of candidate sorting can be integrated. 55% of respondents reported using social media (including LinkedIn) to match candidates with jobs, and 20% reported using search engines or social media to find background information about potential candidates (see Figure 3). These results indicate that indirect use of AI and other ADM is significantly higher than direct use and that these uses are usually tied to a software solution for which some support is available.

There are two further analytical distinctions to be made here. The first is that there is some overlap between the categories of direct and indirect use and that many of the options we offered the survey respondents are not unambiguously one or the other. Without explicitly asking whether the systems in question were developed in-house, the best we can offer is an estimate of the extent to which they fall in one category or the other. The second challenge is that indirect use covers a wide variety of products and services. A further distinction can be made between those uses that are easily accessible to the public (e.g., Google search) and those that were acquired by the company and therefore can be expected to come with support (e.g., Varbi). Another can be made between products that include AI as an additional feature (e.g., Microsoft Word) and those with a feature set largely built around AI models (e.g., LinkedIn Recruiter). The latter is associated with the potential risks of discrimination stemming from how AI models are trained, but with the addition that it is difficult for any external actor to audit proprietary and large-scale systems. Most indirect uses fall somewhere between these extremes.

In the findings detailed above, respondents were quick to highlight when they made direct use of AI or ADM but largely ignored indirect use. Given the format of the questionnaires, it is unclear whether this was out of ignorance or denial, but it is a question deserving of further research.

5.2. The Ubiquity of LinkedIn

LinkedIn, given its ubiquity in recruitment processes (Ajunwa & Greene, 2019; Laukkarinen, 2023; Ruparel et al., 2020), is of particular interest here. As it is not possible to completely differentiate LinkedIn from other recruitment platforms in the questionnaire results (especially as many firms use a mix of recruitment methods), the extent to which AI-powered solutions are used in recruitment processes cannot be measured definitively. Nevertheless, the results indicate that the use of LinkedIn is common, and one interpretation of the results is that around two-thirds of companies that use a recruitment platform use that one in particular. This means that companies' assessment of discrimination risks through the use of AI or other ADM relies heavily on their belief that LinkedIn manages these risks. This suggests that more attention should be given to the role of large-scale digital platforms in addressing discrimination issues in general, and to the measures they take to reduce the risks of discrimination. Although the DSA requires significant transparency from services like LinkedIn, how and if this transparency affects firms that use their services is deserving of further empirical attention.

5.3. Lack of Transparency and (Distributed) Accountability

These concerns around indirect use lead also to a renewed emphasis on the importance of transparency during these processes. Issues related to transparency and accountability are recognised problems, both as governance challenges for AI systems and large digital platforms (e.g., Geng, 2023; Novelli et al., 2023). These are particularly salient here given that the questionnaire data indicate that AI and ADM are used both directly and indirectly, with implications for transparency and accountability. This indirect use may also lead to lower awareness of the risks of discrimination in the use of AI and other ADM in the workplace (e.g., Khatry, 2020). This poses a particular challenge in terms of oversight, where laws in the EU (specifically the DSA) understand large platforms as increasingly responsible for biased content, but not necessarily for the consequences of this content.

Given that both recruitment agencies and major employers underestimate the extent to which they use AI and ADM services while making extensive use of services that have these technologies built in, greater awareness in general is called for. However, opacity in algorithms, and especially complex self-learning AI systems, makes it hard (if not impossible) to assess how and to whom accountability should be allocated (cf. Novelli et al., 2023). While respondents to our questionnaire overwhelmingly thought that the company using the system should be held responsible for bias (84%; see Figure 4), this risks allocating responsibility to actors who do not influence the underlying algorithmic models. This is possible both in direct use where a service has not been developed in-house, and a given when talking about indirect use of services offered by large digital platforms. More transparency would help in offering a more nuanced approach to accountability in what we think of as being distributed rather than shared (Dignum, 2019).

Regulations that have attempted to increase transparency (including the DSA) have faced numerous stumbling blocks. Among them is an inherent conflict of interest between the need for transparency for a supervisory authority—for example, to be able to review how an AI system is developed—and companies' interest in not disclosing how their products work in a competitive market (Larsson & Heintz, 2020). Still, another is the argument that complex algorithms, for instance, large language models (LLMs), are inherently opaque and that it is just not technically feasible to be fully transparent (cf. Peters, 2023). Transparency brings with it the risk of abuse and workarounds—where users learn to subvert responsibility (cf. De Laat, 2018). While this article has focussed on recruitment firms and major employers, transparency alone does not give small employers the resources to either comply with or remedy concerns around bias and risk in AI and ADM use.

Large digital platforms, of which LinkedIn is one, highlight that transparency is important not just at the level of how data are handled and how algorithms draw the conclusions that they do (e.g., Larsson et al., 2021). Rather, the extensive use of such platforms (including unreflexively) adds to a shift in the balance of power in a society, where societal functions are controlled by a few globally active—largely American-anchored—technology conglomerates (Larsson, 2021; van Dijck et al., 2018).

6. Conclusion

This study has aimed to contribute to knowledge about (a) the extent to which AI and ADM are used in recruitment processes, (b) usage perception, and (c) questions around accountability in light of concerns around AI and ADM leading to bias in recruitment. While recruitment companies and major employers overwhelmingly said that their use of ADM/decision support in recruitment was fairly low or very low, in practice all respondents of both categories made use of third-party digital platforms, including social media and recruitment services like LinkedIn, that use AI or ADM to varying degrees. While it is hard to measure, then, the exact extent to which recruitment agencies and major employers use AI or ADM in recruitment (given overlaps and the variety of definitions and degrees of AI use) differs depending on whether one talks about direct use or indirect use.

This extensive indirect use of AI and ADM raises concerns about awareness and reflexivity during recruitment processes. It also has implications for accountability and oversight of the discrimination and bias risks implicit in AI and ADM use. While there have been significant advances made in compelling transparency among, for instance, large digital platforms (through the DSA) and high-risk uses of AI (through the AI Act), this does not necessarily solve the question of who is accountable for bias or discrimination that may result from AI or ADM use in recruitment. While large employers were quick to say that they should themselves be held accountable (84%), transparency would make it easier to decide how, and to whom, accountability and responsibility should be allocated.

Overall, the use of AI and ADM is already well-established in Sweden, and its prevalence is likely to be an exemplar case of what recruitment looks like in the rest of the world. While the exact numbers and platform names might differ, the issues raised here around awareness, indirect use, transparency, and accountability are likely to be generalizable to other contexts. The use of questionnaires here to access busy professionals allowed us to gain a snapshot view of the issues. The exact dividing lines between the (direct and indirect) use of technologies with different features, as well as the motivations and trade-offs that inform AI and ADM use, are beyond the scope of this data. However, the results here suggest that these more granular issues around AI and ADM use in recruitment are deserving of further attention by both researchers and policymakers.

Acknowledgments

The authors would like to extend their gratitude to the Swedish Equality Ombudsman (DO) as well as the Research Institute for Sustainable AI (RISAI) for significant assistance in the execution of this study.

Funding

This study has in part been funded by (a) the Swedish Equality Ombudsman and (b) the Wallenberg AI—Autonomous Systems and Software Program–Humanities and Society (WASP-HS) funded by the Marianne and Marcus Wallenberg Foundation and the Marcus and Amalia Wallenberg Foundation.

Conflict of Interests

The authors declare no conflict of interests.

Data Availability

The data can be made available to bone fide researchers upon request.

References

Ajunwa, I. (2023). *The quantified worker: Law and technology in the modern workplace*. Cambridge University Press.

Ajunwa, I., & Greene, D. (2019). Platforms at work: Automated hiring platforms and other new intermediaries in the organization of work. In S. P. Vallas & A. Kovalainen (Eds.), *Work and labor in the digital age* (Vol. 33, pp. 61–91). Emerald Publishing.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). ACM.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the New Jim code*. Polity Press.

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS Quarterly*, *45*(3), 1433–1450. https://doi.org/10.25300/MISQ/2021/16274

Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In *NIPS'16: Proceedings of the 30th international conference on neural information processing systems* (pp. 4349–4357). Curran Associates Inc.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In X. Alameda-Pineda, M. Redi, E. Celis, N. Sebe, S.-F. Chang (Eds.), *Proceedings of the 1st conference on fairness, accountability and transparency* (pp. 77–91). ACM.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

D'Ignazio, C., & Klein, L. F. (2023). *Data feminism*. MIT Press.

De Laat, P. B. (2018). Algorithmic decision-making based on machine learning from big data: Can transparency restore accountability? *Philosophy & Technology*, *31*(4), 525–541.

Dencik, L., & Stevens, S. (2021). Regimes of justification in the datafied workplace: The case of hiring. *New Media & Society*, *25*(12), 3657–3675. https://doi.org/10.1177/14614448211052893

Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer.

European Commission. (2021). *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (the AI Act) and amending certain Union Legislative Acts* (COM/2021/206 final).

European Commission. (2022). *Digital Services Act* (Regulation (EU) 2022/2065, DSA).

European Commission. (2023). *Artificial intelligence—Questions and answers.*

Ferrara, E. (2023). *Should ChatGPT be biased? Challenges and risks of bias in large language models*. Arxiv. https://arxiv.org/abs/2304.03738

Gasser, U., & Almeida, V. (2017). A layered model for AI governance. *IEEE Internet Computing*, *21*(6), 58–62.

Gaucher, D., Friesen, J., & Kay, A. C. (2011). Evidence that gendered wording in job advertisements exists and sustains gender inequality. *Journal of Personality and Social Psychology*, *101*(1), 109–128.

Geng, Y. (2023). Transparency for what purpose? Designing outcomes-focused transparency tactics for digital platforms. *Policy & Internet*. Advance online publication. https://doi.org/10.1002/poi3.362

Görman, U., & Etikprövningsmyndigheten. (2023). *Vägledning—Om etikprövning av forskning på människor*. Etikprövningsmyndigheten.

Hagendorff, T., & Wezel, K. (2020). 15 challenges for AI: Or what AI (currently) can't do. *AI & Society*, *35*, 355–365.

Ilsøe, A., & Söderqvist, C. F. (2022). Will there be a Nordic model in the platform economy? Evasive and integrative platform strategies in Denmark and Sweden. *Regulation & Governance*, *17*(3), 608–626.

Insight Intelligence. (2022). *The Swedish people and AI: Swedish people's attitudes towards artificial intelligence 2022*.

Jones, P. (2021). *Work without the worker: Labour in the age of platform capitalism*. Verso.

Kaun, A. (2022). Suing the algorithm: The mundanization of automated decision-making in public services through litigation. *Information, Communication & Society*, *25*(14), 2046–2062.

Kaun, A., Lomborg, S., Pentzold, C., Allhutter, D., & Sztandar-Sztanderska, K. (2023). Crosscurrents: Welfare. *Media, Culture & Society*, *45*(4), 877–883. https://doi.org/10.1177/01634437231154777

Kelly-Lyth, A. (2021). Challenging biased hiring algorithms. *Oxford Journal of Legal Studies*, *41*(4), 899–928.

Khatry, S. (2020). Facebook and Pandora's box: How using big data and artificial intelligence in advertising resulted in housing discrimination. *Applied Marketing Analytics*, *6*(1), 37–45.

Kim, K., & Moon, S. I. (2021). When algorithmic transparency failed: Controversies over algorithm-driven content curation in the South Korean digital environment. *American Behavioral Scientist*, *65*(6), 847–862.

Komljenovic, J. (2019). Linkedin, platforming labour, and the new employability mandate for universities. *Globalisation, Societies and Education*, *17*(1), 28–43.

Larsson, S. (2019). The socio-legal relevance of artificial intelligence. *Droit et Société*, *103*(3), 573–593.

Larsson, S. (2021). Putting trust into antitrust? Competition policy and data-driven platforms. *European Journal of Communication*, *36*(4), 391–403.

Larsson, S., Haresamudram, K., Högberg, C., Lao, Y., Nyström, A., Söderlund, K., & Heintz, F. (2023a). Four facets of AI transparency. In S. Lindgren (Ed.), *Handbook of critical studies in artificial intelligence* (pp. 445–455). Edward Elgar Publishing.

Larsson, S., & Heintz, F. (2020). Transparency in artificial intelligence. *Internet Policy Review*, *9*(2). https://doi.org/10.14763/2020.2.1469

Larsson, S., Jensen-Urstad, A., & Heintz, F. (2021). Notified but unaware: Third-party tracking online. *Critical Analysis of Law*, *8*(1), 101–120.

Larsson, S., Liinason, M., Tanqueray, L., & Castellano, G. (2023b). Towards a socio-legal robotics: A theoretical framework on norms and adaptive technologies. *International Journal of Social Robotics*, *15*, 1755–1768. https://doi.org/10.1007/s12369-023-01042-9

Laukkarinen, M. (2023). Social media as a place to see and be seen: Exploring factors affecting job attainment via social media. *The Information Society*, *39*(4), 199–212.

Leavy, S. (2018). Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning. In *Proceedings of the 1st international workshop on gender equality in software engineering* (pp. 14–16). ACM.

LinkedIn. (2018). *An introduction to AI at LinkedIn*. https://engineering.linkedin.com/blog/2018/10/an-introduction-to-ai-at-linkedin

Lomborg, S. (2022). Everyday AI at work: Self-tracking and automated communication for smart work. In S. Pink, M. Berg, D. Lupton, & M. Ruckenstein (Eds.), *Everyday automation: Experiencing and anticipating emerging technologies* (pp. 126–139). Routledge.

Morondo Taramundi, D. (2022). Discrimination by machine-based decisions: Inputs and limits of anti-discrimination law. In B. Custers & E. Fosch-Villaronga (Eds.), *Law and artificial intelligence: Regulating AI and applying AI in legal practice* (pp. 73–85). TMC Asser Press.

Myers West, S., Whittaker, M., & Crawford, K. (2019). *Discriminating systems: Gender, race, and power in AI*. AI Now Institute.

Newlands, G. (2022). Anthropotropism: Searching for recognition in the Scandinavian gig economy. *Sociology*, *56*(5), 821–838. https://doi.org/10.1177/00380385211063362

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

Novelli, C., Taddeo, M., & Floridi, L. (2023). Accountability in artificial intelligence: What it is and how it works. *AI & Society*. https://doi.org/10.1007/s00146-023-01635-y

Pasquale, F. (2020). *New laws of robotics: defending human expertise in the age of AI*. The Belknap Press of Harvard University Press.

Peters, U. (2023). Explainable AI lacks regulative reasons: Why AI and human decision-making are not equally opaque. *AI and Ethics*, *3*(3), 963–974.

Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2016). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, *20*(1), 293–310. https://doi.org/10.1177/1461444816661553

Roehl, U. B. U. (2022). Understanding automated decision-making in the public sector: A classification of automated, administrative decision-making. In G. Juell-Skielse, I. Lindgren, & M. Åkesson (Eds.), *Service automation in the public sector. Concepts, empirical examples and challenges* (pp. 35–63). Springer Nature.

Ruparel, N., Dhir, A., Tandon, A., Kaur, P., & Islam, J. U. (2020). The influence of online professional social media in human resource management: A systematic literature review. *Technology in Society*, *63*. https://doi.org/10.1016/j.techsoc.2020.101335

Russell, S. (2019). *Human compatible: AI and the problem of control*. Penguin Books.

Sánchez-Monedero, J., Dencik, L., & Edwards, L. (2020). What does it mean to "solve" the problem of discrimination in hiring? Social, technical and legal perspectives from the UK on automated hiring systems. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 458–468). ACM.

Statistics Sweden. (2023). *Antalet anställda i snabbväxande storföretag kommer att öka*. https://www.scb.se/hitta-statistik/temaomraden/sveriges-ekonomi/fordjupningsartiklar_Sveriges_ekonomi/antalet-anstallda-i-snabbvaxande-storforetag-kommer-att-oka

Söderlund, K. & Engström, E. & Haresamudram, K. & Larsson, S., & Strimling, P. (2024). Regulating high-reach AI: On transparency directions in the Digital Services Act. *Internet Policy Review*, *13*(1). https://doi.org/10.14763/2024.1.1746

Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society*, *40*(2), 137–157.

Thylstrup, N. B. (2018). *The politics of mass digitization*. MIT Press.

van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197–208.

van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connective world.* Oxford University Press.

Veale, M., & Zuiderveen Borgesius, F. (2021). Demystifying the Draft EU Artificial Intelligence Act: Analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International*, *22*(4), 97–112.

Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review*, *41*. https://doi.org/10.1016/j.clsr.2021.105567

Woodcock, J., & Graham, M. (2020). *The gig economy: A critical introduction*. Polity Press.

Zhang, L., & Yencha, C. (2022). Examining perceptions towards hiring algorithms. *Technology in Society*, *68*. https://doi.org/10.1016/j.techsoc.2021.101848

About the Authors

Stefan Larsson is a senior lecturer and associate professor in technology and social change at Lund University, Sweden, Department of Technology and Society. He is a lawyer and socio-legal researcher who holds a PhD in sociology of law as well as a PhD in spatial planning. He leads a multidisciplinary research group that focuses on issues of trust and transparency and the socio-legal impact of autonomous and AI-driven technologies in various domains, such as consumer markets, the public sector, and in health and social robotics.

James Merricks White is a postdoctoral researcher in the Department of Technology and Society, Lund University, where he is studying trust and trustworthiness in AI consumer markets. His research interests include expertise, knowledge/ignorance, and standardisation, with a focus on the International Organization for Standardization (ISO) and the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES). He has recently been awarded a VR grant (2023-01549) to study risk governance mechanisms in the AI Act.

Claire Ingram Bogusz is an associate professor (docent) in the Department of Informatics and Media, with a background in Law, philosophy, and economics. She conducts research into how code-based technologies (e.g., code itself, digital platforms and infrastructures, learning machines, and digital automation) change how we work, how organising occurs, and what these things mean for us as individuals and as members of increasingly polarised and unequal societies.