Futures 166 (2025) 103538

Contents lists available at ScienceDirect

Futures

journal homepage: www.elsevier.com/locate/futures

On the present-future impact of AI technologies on personnel selection and the exponential increase in meta-algorithmic judgments

Lambros Roumbanis [a,b]

[a] Institute for Futures Studies, Sweden
[b] Department of Sociology, Örebro University, Sweden

ARTICLE INFO

Keywords:
Artificial intelligence (AI)
Human judgment and decision-making
Hiring and personnel selection
Expert recruiters
Expectations and imagined futures
Sociology of algorithms

ABSTRACT

The ongoing implementation of new AI technologies in both public and private organizations has already had a great impact on professional judgment and decision-making processes. For many experts working with people's applications tied to, e.g., welfare services, employment, school admission, parole, and insurance, algorithm-based decision assistance has already changed the way they perform their assessments. Yet we are still in what seems to be the early stages of this socio-technological transformation. Based on interviews with twenty-four expert recruiters, the purpose of the study is to explore their professional views and expectations concerning the present-future impact of AI on personnel selection. The study seeks to shed new light on the exponential increase in meta-algorithmic judgments, a concept that pinpoints the fact that many experts today must handle algorithmic assessments before moving forward in the decision-making process. What do the respondents think has motivated this socio-technological change, and what do they expect will happen to the role of human judgment in the future? The analysis suggests that the seemingly inevitable existence of human biases and the lack of efficiency in traditional recruitment are, according to all respondents, the main reasons why AI should be used, given its promise to improve both fairness and time management. While AI recruitment systems will most likely become increasingly sophisticated, most respondents thought that these systems would nevertheless never be able to fully replace human judgment. However, a handful of respondents deviated from this majority view by instead expecting that AI would take over the whole process in the not-too-distant future.

1. Introduction

"*Who* programs *what*? *What* programs *who*? Does the *who* program the *what* through self-programing?"

– Bernard Stiegler

In the early 1980s, St. George's Hospital Medical School in London began using a computer program to screen student applications for admission (see, e.g., Iganski & Mason, 2018; Lowry & MacPherson, 1988). The project was initiated by the then Vice-Dean Geoffrey Franglen, himself an admissions assessor who wanted to improve the process by automating the screening part. By implementing an algorithmic system, his main concern was to increase efficiency, but he also hoped doing so would increase fairness and consistency.

E-mail address: lambros.roumbanis@iffs.se.

https://doi.org/10.1016/j.futures.2025.103538
Received 12 June 2023; Received in revised form 1 December 2024; Accepted 4 January 2025
Available online 6 January 2025

About three-quarters of the approximately 2000 annual applicants had to be weeded out, and those who managed to pass the screening threshold were invited for an interview. Around 70 percent of those who performed the interview were finally offered enrollment in the educational program. Needless to say, the screening was a decisive gatekeeping mechanism, and replacing human judgment with a properly designed algorithmic system would presumably lead to all applicants being subject to the same assessment. Thus, Franglen created his program after a careful analysis of the entire process and double-tested it with the work done by human assessors. The algorithms performed well, giving a 90–95 percent correlation with the scores of the admissions staff. However, after a few years with the new automated screening process, two senior lecturers at St. George's became seriously concerned by the lack of diversity among the newly accepted student cohorts, and they decided to report their observation to the board. A special commission found that the system unintendedly discriminated against women and people with non-European-sounding names. In retrospect, the algorithmic system only did what it had been programed to do – mimic the previous assessment records of real human assessors and score student admissions according to predetermined assessment criteria. And yet, the algorithms accidentally amplified already existing patterns of social exclusion and discrimination.

Over forty years have passed, and a range of new deep learning AI technologies are becoming an integral part of evaluation and decision-making processes. Today, there is a strong collective expectation among many stakeholders and policymakers that this ongoing techno-scientific development will help to radically improve human societies in the future. Governments all over the world have declared their national AI strategies in policy documents, and they have invested innumerable billions in R&D annually. Private companies, tech industries, and universities have all played a pivotal role in innovating and promoting the latest generation of AI technologies (Bareis & Katzenbach, 2022; Hoff, 2023; Schiølin, 2019). Meanwhile, ordinary people and professionals have become more used to being supported by algorithmic recommendations, which they must deal with before making decisions. These new synthesized decision-making relations between algorithms and human judgment can be conceptualized as *meta-algorithmic judgments*, which refers to the 'procedural relationship in time between the evaluative output produced by algorithms and the judgment that must be made by a human user in the next step' (Bursell & Roumbanis, 2024, p.3). This conceptual approach thus highlights an important socio-technical dimension, that is, the creation of 'conjoined agencies' between human and AI in everyday life (e.g., Jarrahi et al., 2022; Murray et al., 2021; van den Broek et al., 2021). Yet the notion of meta-algorithmic judgment also adds to previous discussions concerning the potential dilemmas that occur when people either feel perhaps an unnecessary aversion to algorithms (Burton et al., 2020; Dietvorst et al., 2015; Prahl & Van Swol, 2017) or for some reason are too reliant on them (Logg et al., 2019; Parasuraman & Manzey, 2010; Wagner et al., 2018).

But what exactly motivates and justifies the vast implementation of AI we are currently witnessing? And what are the arguments against this development? To be sure, behavioral and social science scholars have played an active role in highlighting the benefits of AI-powered decision-making in, for instance, healthcare, child custody, asylum, hiring, bail decisions, and resource allocation (Jaton, 2021; Kahneman et al., 2021; Logg et al., 2019; Raghavan et al., 2019). One issue that has been emphasized in particular is how AI can help eliminate the harmful impact of so-called 'implicit biases,' thus improving both general fairness and public health (Greenwald et al., 2022; Lin, Hung & Huang, 2021). These are some of the arguments that have been presented to justify the development and implementation of AI in social organizations. But like many other emerging technologies with potentially disruptive powers, AI comes not only with great hopes, but also with deep ethical concerns and anxieties. While much of the discussion about AI is forward looking, a growing number of social science and legal scholars have expressed strong criticism of how machine-learning algorithms are currently being used to determine people's life chances, pointing to the problems associated with giving too much power and responsibility to these technologies (Ajunwa, 2022; Brayne, 2020; Howard & Borenstein, 2018; Kiviat, 2023; Pasquale, 2020). A constant matter of concern for many of these scholars has revolved around the problem of 'algorithmic bias,' namely how biased datasets or biased algorithms may cause an AI to discriminate against socially salient groups or members of these groups (Bursell & Roumbanis, 2024; Köchling & Wehner, 2020; Seppälä and Małecka, 2024).

The ongoing algorithmic transformation of human societies, which has in different ways changed how people interact and make decisions, has also raised questions concerning its potentially alienating effects (Burrell & Fourcade, 2021; Hou, 2021). Hence, AI technology seems to have both pros and cons (like nuclear power, which produces both electricity and toxic waste), and the uncertainty surrounding the future development of AI leaves many people alternating between optimism and pessimism. How, then, can we further explore these critical issues? What appears to be a crucial task from a sociological perspective is to anchor previous research and forecasts concerning how AI can transform society in the expectations of either the professionals responsible for creating AI software (computer engineers, coders, AI researchers) or the groups of professionals responsible for using the systems in their daily practices (e.g., social workers, police officers, HR managers). Listening to what they have to say about AI can enable an in-depth examination of *technics*, 'apprehended as the horizon of all possibility to come and of all possibility of a future' (Stiegler, 1998, p. ix).

In the present study, my main objective is to explore the present-future relationship between human judgment and AI, and to search for answers to the question of how this ongoing technological transformation of social life can change the way organizations will select people for job opportunities in the future. I will accomplish this aim by exploring the views and expectations of a group of professionals whose use of AI-powered assessment systems has increased significantly during the past few years, namely *expert recruiters*. Based on 24 semi-structured interviews with recruiters in key positions, working in hiring agencies, HR departments at large municipalities or big companies in Sweden, the purpose is to investigate what they express when reflecting on the present-future impact of AI on their profession. The structure of the paper is as follows: (i) background, theoretical contextualization, and a basic typology of meta-algorithmic judgments; (ii) methods and data collection; (iii) empirical results and analysis; and (iv) concluding discussion. A central aim of this inquiry is to contribute to the sociology of algorithms and the general understanding of human-technology relations. The two overall questions posed in the study are: *What motivates the use of AI in personnel selection, and how will more advanced AI potentially change the role of human judgment in future hiring*?

2. Artificial intelligence and human judgment

2.1. From human bias to algorithmic bias: is unbiased decision-making possible?

The accelerating progress in computer engineering and research on deep neural networks, natural language processing (NLP), and machine learning algorithms have facilitated new forms of common learning process in human-AI interactions. Yet one of the greatest challenges today seems to revolve around how human intelligence and artificial intelligence can be successfully synthesized. Compatibility and trust are of paramount importance to making that happen. In a recent paper, Jarrahi and colleagues (2022) emphasized the theoretical value of comparing and contrasting AI with human intelligence by distinguishing their unique capabilities in relation to each other. For example, at present a generative AI model that is built on NLP can conduct structured interviews with job candidates and then score their performances in line with pre-defined templates, but do such models really understand what they are evaluating? To be sure, there is currently a rather heated debate within the AI research community concerning whether NLP systems like ChatGPT and BERT actually can be said to 'understand' the meaning they process like humans do, or whether it is more appropriate to talk about a non-human form of comprehension (e.g., Arkoudas, 2023; Bender et al., 2021; Fazi, 2021; Mitchell & Krakauer, 2023). In any case, there is an important contrast here between humans and AI that needs to be taken into account. Humans have the unique capability to form a judgment based on deep reflection, contrafactual thinking, deliberation, and moral sensibility, whereas AI systems are still bound to their capacity to calculate things, and this difference is fundamental. Although he coined the notion of 'computational reflection' decades ago, Cantwell Smith (2019), p. xiii) recently declared:

> "Neither deep learning, nor other forms of second-wave AI, nor any proposals yet advanced for third-wave, will lead to genuine intelligence. Systems currently being imagined will achieve formidable reckoning prowess, but human-level intelligence and judgment, honed over millennia, is of a different order. It requires 'getting up out' of internal representations and being committed to the world *as world*, in all its unutterable richness."

Since at least Aristotle (350BC/(1955), philosophers have been treating the faculty of judgment as a very unique and special aspect of human intelligence. An expert judgment implies a deep sensibility that can be used to discern a certain quality or phenomenon, thus a capability that all knowledge and science fundamentally depend upon. However, research from cognitive psychology has later depicted human judgment as inherently biased when it is guided by intuition or 'gut feelings,' resulting in human actors seeming to be hopelessly irrational in many situations. Even expertise and professional judgments have been shown to be biased, including the conclusions of statisticians, academic evaluators, or court judges (Roumbanis, 2017). Moreover, sociologists have shown how hiring experts are heavily influenced by their emotions and biases in personnel selection (Pedulla, 2020; Rivera, 2015). These insights have planted profound doubt at the very core of public institutions and private companies, suggesting that crucial decisions may be inevitably flawed by the heuristic principles that guide human judgment.

Even if many of us like to believe we are unprejudiced and open-minded when we encounter other people in our everyday life, this belief was called into serious question when Greenwald and Banaji (1995) presented their 'Implicit Association Test' (IAT), showing that our social cognition is inevitably formed by associations we have habitualized since early childhood. These associations guide our preferences and actions in all kinds of social situations, almost like knee-jerk responses. Indeed, this discovery is in close accordance with certain branches of modern cognitive sociology (Bursell, 2024; Lizardo, 2022; Olsson, 2024; Shepherd, 2011).

Implicit biases are, in other words, activated at a preconscious level in the human mind and may evoke discriminatory behavior in all people without them noticing it. Although these scientific results are far from undisputed among contemporary researchers and experts,[1]

what is important to acknowledge is that these ideas have been highly influential within HR communities in shaping expert recruiters' self-understanding. Again, this is where decision support from AI becomes important, not least in light of the claims offered by many of today's AI recruitment companies regarding the debiasing effects their technologies can provide (Drage & Mackereth, 2022; Seppälä and Małecka, 2024). The problem, however, is that machine-learning algorithms are far from immune to implicit biases. There are already several known cases in which specialized AI applications have found and then operated based on patterns in large datasets 'that reflect implicit biases and, in so doing, emphasize and reinforce these biases' (Howard & Borenstein, 2018, p.1522). And even if more advanced AIs could be created in the future, seemingly capable of preventing human biases from permeating the decision-making processes, the question is: Would human supervisors be able to monitor the accuracy and reliability of these systems in practice? As computer scientist Yampolskiy (2024, p.6) recently noted, the monitorability of today's increasingly complex and creative machine-learning systems already presents a significant challenge in the pursuit of trustworthy AI because, 'emergent properties cannot be pre-detected, as they arise from complex interactions that are not explicitly encoded in the AI system's design.'

2.2. On contemporary AI hiring technologies

Today there is a plethora of different AI tools available on the market that cover almost every aspect of human resource management, talent acquisition and recruitment. Some of these hiring technologies use more traditional rule-based algorithms in combination with integrated psychometric assessment methods to screen job candidates' work-related competencies and problem-solving

[1] There has been an ongoing debate among social psychologists and sociologists regarding the very nature of implicit bias and whether such biases can be reduced through anti-bias training (see, e.g., Bursell, 2024).

behavior (Bursell & Roumbanis, 2024; Köchling & Wehner, 2020). Although these kinds of algorithmic hiring technologies have existed for many years, they have nonetheless continued to be refined, making them increasingly sophisticated for use in personnel selection. Some of these rule-based AIs are designed as questionnaires using the 'Five Factor Model' to measure candidates' suitability and predict behaviors critical for working life.

More recent innovations in the area of candidate screening are CV screening tools and *interview chatbots* (conversational AI). These technologies are based on natural language processing, and some of the more advanced chatbots utilize a deep learning model called BERT (Bidirectional Encoder Representations from Transformers) that has been pre-trained on a vast array of texts to 'understand' the subtilities of human language. An interview chatbot built on a BERT model can be trained to generate rich contextual word embeddings to map both familiar and novel concepts onto a shared semantic space; the chatbot basically learns to represent text using 'self-supervised learning' (Rogers et al., 2020; see also Mitchell & Krakauer, 2023). This approach enables the machine-learning algorithms to recognize and classify text into new themes that were not part of its initial training data, the goal being to 'evaluate,' for example, which of the candidates are the most qualified for a certain position, depending on how well their answers match the pre-defined criteria and parameters for a certain job profile. Another type of AI chatbot is the more famous ChatGPT ('generative pre-trained transformer'), which is based on a combination of supervised learning and reinforcement learning from human feedback (Arkoudas, 2023).

Perhaps one of the most controversial AI software systems in the hiring industries to date are those that combine facial/body recognition technology, psychometric and emotional assessment, and speech recognition technology to analyze job candidates. Today, several vendors (e.g., *HireVue*, *Retorio*, *myInterview*) offer online platforms on which candidates can conduct fully automated video interviews, which are claimed to be based on algorithmic models that are 'blind' to race/ethnicity, age, and gender, thus helping organizations make unbiased hiring decisions and build a more diversified workforce. In a recent article, Drage and Mackereth (2022, p.16) critically noted that this AI-powered hiring tool is constructed to:

> "…analyze the minutiae of a candidate's speech and bodily movements in order to measure how closely they respond to the employer's 'ideal employee.' In the case of facial recognition software, the …head movement, expression, and intonation is tracked and interrogated to gauge whether or not a candidate is the 'right fit.' Meanwhile, speech recognition software analyzes vocabulary and word choice to discern whether or not a candidate is truly 'passionate' about the job and whether their enthusiasm is genuine."

Hence, this AI registers people's speech, gestures, and 'micro-expressions' during their online presence inside the hiring platform, thus constructing new actionable correlations and classifications, which will ultimately affect whether a person will be hired. Yet efforts to eradicate differences associated with people's gender and race/ethnicity sometimes reveal a prevailing misunderstanding of what gender and race/ethnicity are 'and whether they can be defined as isolatable and removable attributes within hiring' (Drage & Mackereth, 2022, p.2; see also Ajunwa, 2022). How will the algorithms manage to analyze whether a person is the 'right fit' or 'truly passionate' using such a stripped-down approach? Another issue that can be raised is how a video hiring AI can evaluate the facial and body movements of persons with different disabilities in an unbiased manner. For example, as Buyl et al. (2022, p.1078) underscored, 'candidates with autism may avoid eye gaze during an interview, which may be misinterpreted as a lack of confidence in their abilities.' There exists, in other words, an important critical debate on the widespread claims about 'unbiased recruitment' related to the current AI hiring tools (Köchling & Wehner, 2020; Howard & Borenstein, 2018).

2.3. A typology of meta-algorithmic judgments

Looking back historically, a feeling of discomfort with handing over some of the control to algorithm-based technologies has not infrequently emerged among professionals. For example, Lauer (2017) highlighted the introduction of statistical credit scoring in the US during the 1960s. Despite some of the obvious benefits of using risk-calculating technologies, credit managers at the time refused to embrace it, dismissing it as an inadequate substitute for human judgment. As Lauer underscored, creditworthiness was believed to be rooted in the human character, an elusive moral trait that must be judged by skilled professionals through face-to-face interviews. Hence, human intuition and evaluative sensibility were deemed fundamental to making appropriate decisions. But since then, of course, a great deal has happened, and the balance between humans and algorithms has changed quite dramatically.

However, mismatches regularly occur between AI advice and the way humans consider such advice when making decisions. In a recent study, Bursell and Roumbanis (2024) compared the outcomes of traditional versus algorithm-based recruitment processes during implementation of a new AI system at one of Sweden's largest food retail companies. By drawing on two large datasets of employment records and job applications during the transition period 2019–2020, the study showed that the algorithm-based recruitment process decreased the levels of diversity. But in contrast to what is often assumed, this was not because the algorithmic system was biased in any observable way, but because of what was identified as an unintended human-AI interaction effect. Thus, bias occurred during the different phases of the hiring processes, showing that candidates with non-European surnames were disfavored, despite having equivalent test scores. Bursell and Roumbanis (2024) introduced the concept of *meta-algorithmic judgment* to pinpoint how the recruiting managers had to adapt their hiring routines to a new situation in which algorithms gave them recommendations on job candidates, a task they themselves previously had control over. The recruiters' discretionary power was key in the selection process, thus putting the issue of compliance/noncompliance at the center of the analytical framework. It seemed quite clear, at the aggregated level, that the managers often followed their own 'gut feelings' and preferences when deciding which of the shortlisted candidates to invite for a personal interview, and then later, which to hire. Hence, by deviating from or even ignoring the meritocratic structure of the algorithmic advice, the recruiting managers' meta-algorithmic judgments caused the levels of diversity to

decrease within this very large multi-site organization (Bursell & Roumbanis, 2024, p.14).

In what follows, I will present three different types of meta-algorithmic judgments, all of which presuppose different degrees of AI development and different kinds of synthesis between human-AI evaluations. First, we have *discretionary* Type 1 judgment, where the expert recruiter is sanctioned to evaluate and make sense of the algorithms' assessments of the candidates. At present, this is the most common type of meta-algorithmic judgment in AI-powered recruitment. Type 1 can be divided into compliancy/non-compliancy approaches based on the degree to which the human recruiter follows the algorithmic recommendations when selecting candidates for further assessments (Bursell & Roumbanis, 2024). A common learning process between human and AI can be seen as a constant work in progress in organizations. The final hiring decisions will always be made by the expert recruiter. Then there is the *monitoring* Type 2 judgment, where the recruiter has the role of a control agent who overviews the constant flows of job candidates in the AI assessment system, intervening only if deemed necessary or for a supervised learning process. This future synthesis between AI and human intelligence would put meta-algorithmic judgment in a different mode. Here, most assessments and hiring decisions would be fully automated, owing to the deep sensibility, efficiency and precision of the AI system in finding and selecting new candidates for employment. Type 2 would mostly consist of following predetermined routines for monitoring the hiring processes and evaluating a number of integrated control mechanisms. And finally, we have the *disengaged* Type 3 judgment, which refers to a future scenario in which AI has become fully autonomous and makes its own hiring decisions without any direct human involvement. The AI will constantly engage in self-training and improve itself by learning from various sources of information. However, Type 3 implies the very end of human judgment in the hiring process, at least from an organizational perspective. *Disengaged* Type 3 meta-algorithmic judgment can only manifest its critical capacity in hindsight, if something has clearly gone totally wrong, and the AI system needs to be thoroughly scrutinized, or even shut down.

The task of creating genuinely successful human-AI symbiosis for augmented decision-making is far from a trivial matter. The challenge of making humans trust and understand the algorithms would seem to be pivotal. As Burton and colleagues (2019) highlighted, one way to avoid so-called 'algorithm aversion' among professionals and lay people alike may partly depend on the ability to bridge human intuition and AI rationalities. But it is also important that humans not too easily accept algorithmic recommendations and begin taking them for granted, because it is clear that overreliance on AI systems may cause problems if they later turn out to be biased (e.g., Wagner et al., 2018).

3. Methods and data collection

Interviewing people can be an effective way not only of capturing people's own personal views and feelings, but also of grasping more general trends and common perceptions within a wider community. The present study uses a sociological approach that focuses on socio-technical expectations and 'imagined futures' (Beckert, 2016; Konrad, 2006; van Lente, 2012; Oomen et al., 2022). The study is part of a larger ongoing research project looking at the organizational implementation of AI in job recruitment and human-AI interaction effects, and the aim is to complement quantitative and register data studies with qualitative in-depth knowledge.

For that specific purpose, I have conducted 24 semi-structured interviews with expert recruiters (14 women and 10 men). At the time of the interviews, 9 of the respondents worked at recruitment agencies, 9 worked at the HR departments of larger companies, and 6 worked at HR departments in municipalities. One specific criterion that guided my selection of respondents was that they had to have experience of working with AI-based hiring tools (rule-based and/or machine-learning models). After an extensive search for relevant interviewees on organizational webpages and HR forums, I sent emails to 40 expert recruiters, of whom 19 agreed to participate in my study. The same invitation was later sent to another 12 recruiters, which resulted in 5 more interviewees. This sample size turned out to be sufficient for the aim of my study, given the fact that I interviewed a relatively homogenous study population and focused on a comparatively narrowly defined research topic (Hennink & Kaiser, 2022). The interviews were conducted during 2021–22, which initially involved a practical challenge in the form of the COVID-19 pandemic. The Swedish Government was one of very few in the world to decide not to impose a lock-down, however, new restrictions caused many people, including many expert recruiters, to work from their homes to a greater extent than before. I therefore decided to conduct all the interviews online via Zoom. This proved to be a good and efficient solution, as all respondents were used to having meetings on different video communication systems. My own sense from conducting these interviews online was that nothing essential was lost in the communication process (of course, I cannot be completely sure of this). The interviews were between 45 and 60 minutes in duration and were conducted in Swedish. The interview guide targeted the broad topic of hiring, expert judgment and AI, and covered not only the respondents' own personal experiences, but also more general issues and challenges tied to the HR profession. Several of the questions I posed to the respondents related directly to their views on AI and the present-future role of human judgment in the hiring process. All interviews were recorded with the permission of the respondents and thereafter transcribed. The study has been theoretically motivated from the very outset, the intention being to explore and develop the concept of meta-algorithmic judgment by drawing on qualitative material. Given the relatively limited number of interviews I conducted, I preferred to work in a more traditional way, printing the transcriptions and writing codes directly in the margins (Deterding & Waters, 2021). This is a more time-consuming approach, but in my view it resulted in a less mechanical construction of themes than what occurs when using a computer program.

After having carefully read through the material several times, I coded it in line with a modified thematical approach (e.g., Boyatzis, 1998; Byrne, 2022; Christou, 2023), first by focusing on broader topics (organizational change; role of human expertise; experience of working with AI; legitimacy and responsibility). Then in a later round, I made a more detailed coding by focusing in particular on the beliefs and emotions the respondents expressed when reflecting on both traditional and AI-based recruitment (hopes/fears; trust/distrust; promises/pitfalls, etc.). In line with the different codes, I then tried to identify a few main themes and to carve out a basic typology of meta-algorithmic judgment. The first theme I will present below was constructed based on what the

recruiters felt legitimizes and motivates the implementation of AI, which they generally tied to the existence of human bias and the lack of efficiency in traditional hiring (Section 4.1). The second theme sheds light on the factors slowing down the technical transformation of organizations, the main reason many of the respondents mentioned being fear and lack of understanding of what AI is (Section 4.2). In the third theme (Section 4.3), I interpret the meaning of meta-algorithmic judgment at present ('Type 1') by emphasizing the role of discretion and the issue of trust/distrust in human-AI relations. Finally, the fourth theme (Section 4.4) focuses on the recruiters' expectations about future AI technologies and their impact on human judgment. This last theme will be used to analyze the two future-oriented meta-algorithmic judgments 'Type 2' and 'Type 3,' which I constructed based on the recruiters' 'imagined futures' and ideas about the different levels of human involvement in future AI recruitment (i.e., discretion, monitoring, disengaged).

4. Results and analysis

4.1. What motivates and justifies the implementation of AI in hiring?

When I asked the recruiters what they believed were the most prominent problems with traditional hiring, most emphasized two things in particular: *lack of efficiency* and *human biases*. Traditional hiring involves designing evaluation criteria, reading CVs and personal letters, ranking candidates, conducting interviews, and checking references. In Sweden, as in many other countries, recruiting mangers and expert recruiters often have to devote a great deal of time and energy to screening through hundreds of applications in order to select a handful of candidates for further evaluations and interviews. This is an immensely inefficient assessment procedure. Still, cognitive biases and 'gut feelings' (Pedulla, 2020; Rivera, 2015) were deemed by every single respondent to constitute the greatest problem in traditional hiring, because they result in many good and valuable talents being unintendedly disfavored or discriminated against. One 60-year-old female recruiter with nearly 30 years of experience working with hiring processes and HR management said:

> "We humans are hopeless when it comes to discrimination. In our profession as recruiters, we often attend different workshops and on-the-job training, because it's very, very important to follow the latest research and trends in HR. And yet, despite all this knowledge, we still be fall back on our implicit biases, …it really doesn't matter how hard I try to avoid them. You will always be affected by so many different things, and form opinions about other people in milliseconds."

This echoes much of the research on implicit biases in social psychology that I mentioned earlier (Greenwald & Banaji, 1995; Olsson, 2024; Shepherd, 2011), and it highlights the real dilemma that many professionals must deal with in their daily practices, namely the feeling that their own trained judgment might be influenced by both irrelevant and discriminatory factors. One 45-year-old male recruiter put it like this: 'Bias is a well-known problem in our profession. There are biases about gender, ethnicity, age, language use, physical appearance, etc., and you're not conscious of them, and that's a general flaw, …It's very challenging to do traditional hiring with accuracy.'

According to some respondents, the problem of human bias is the reason why the entire HR profession must undergo a technological transformation, because that would ultimately enhance meritocracy and fairness on the labor market. Yet no one thought that humans should be completely replaced by autonomous AI systems. Instead, most recruiters felt that enhanced human-AI synthesis was pivotal for establishing an augmented decision-making process. As one female recruiter around 50 years of age explained:

> "For me, using AI is first and foremost about creating an unbiased process. That's extremely important to me. But I would also like to underscore, that this AI cannot replace the craftmanship of a human recruiter. It may be possible for such techniques to be developed in the future, but we are certainly not there yet. I really think this has to do with the issue of soft skills. How can you ensure that an AI can 'sense' the soft skills of the candidates in the same way humans can? I believe there are a lot of challenges, …I mean, all the qualitative aspects that we humans bring into the process with our ability to form a judgment about another person and social contexts, there is a kind of 'consequence thinking' involved in that. We've discussed this issue plenty of times at our company, whether it will be possible to develop deeper sensitivity in an AI robot."

Most recruiters seemed to feel that the real challenge was how to create the best possible symbiosis between humans and AI, that is, how to pave the way for new forms of 'hybrid intelligences' (Jarrahi et al., 2022). One solution is to give the human evaluator a new place in the hiring process. To be sure, humans are doing preparatory HR work before the formal process starts, but they do not necessarily have to evaluate any candidates before the interview stage. If the algorithms can screen the entire pool of candidates, then human judgment can enter at a later stage instead, that is, *after* the algorithm has presented its recommendations (Bursell & Roumbanis, 2024). At that stage, according to one of the respondents, human judgment will be less harmful:

> "…because it will always do harm, to some extent. That's unavoidable. You have to move subjectivity to a later stage in the process, then it doesn't matter so much if you choose candidate A or B, both will be good anyway. Then it's okay to use your professional gut feelings."

Some of the recruiters did express ambivalence regarding the question of whether AI could become as sensitive as a human being. However, a few respondents speculated about the possibility of a future AI that could imitate human intelligence and develop a deeper human-like sensitivity. Still most were of the firm opinion that humans can never really be replaced by AI. The general opinion was that the complete automation of personnel selection would entail the risk of de-humanizing effects. However, one 55-year-old male recruiter with both an engineering and an HR degree offered a different opinion: 'I imagine that AI in the future will be fully capable of capturing the soft skills and social qualities of the candidates. I don't think it's impossible to make algorithms learn to be much more

sensitive… It isn't really that complex, in my view.'

4.2. What is slowing down socio-technical development?

One common impression today, in Sweden and elsewhere, is that we are living at a time when AI development is extremely rapid (e.g., ChatGPT, self-driving cars, AI-generated art). If we turn to the recruiters' everyday life, then their experiences were quite mixed regarding the pace at which AI is transforming the profession. For example, some respondents stated that the progress was too slow (at least within their organizations), whereas others felt many new things related to AI were happening. Most of them could confirm that there was indeed a great deal of enthusiasm and curiosity on the part of many expert recruiters today, but also that many actors were afraid of taking the first step and were waiting for other actors to go first. One 45-year-old male recruiter mentioned fear as a potential obstacle that could have negative effects on organizations' decisions to invest in new AI technologies.

> "It's always scary when something is new, especially if you don't really understand it, because I think that resistance against AI often has to do with ignorance. People don't understand what an algorithm does, or that it's programed on probabilities, … different criteria and preferences, or whatever is being used when a system is designed. I believe that this fear that occurs, …yes, I believe that many people are afraid, I believe, well, let me put it more diplomatically, that you have to respect this fear."

According to this respondent, top managers and HR professionals still seem to fear that AI may lead to something bad, which generates 'algorithm aversion' (Burton et al., 2020; Dietvorst et al., 2015; Prahl & Van Swol, 2017). However, this fear could be expressed in different ways, and not necessarily entail any aversion to the algorithms themselves, but more to the managerial motivations for using them in personnel selection. For example, some respondents enthusiastically embraced the current development of new AI hiring tools, but at the same time expressed strong concerns. One respondent reflected on a particular type of fear that concerned her:

> "I'm *not* afraid that AI will conquer the world, I'm not worried about all the horror scenarios, because there will always be humans governing and controlling the technologies. I cannot really see how, well, …I might be too naïve, but I think new AI technologies are really going to help us in the future, and that we're going to witness great progress […] But there are also many challenges with the use of AI, and I'm very afraid that money and efficiency will impact it too much. There are so many different AI tools on the market today, and many of these tools produce speedy evaluations, and some of these tools are scoring people on their facial expressions, etc. In some states in the US, these techniques have been prohibited, thank God. But there is really no solid and independent research being done on how these systems work. When we developed our own AI [an interview chatbot], we decided very early on that we would never become a new 'Amazon case.' But I still fear that other organizations will use cheaper, faster, and worse AI tools that will do much more harm than good."

This type of fear, namely that money and efficiency will become the ultimate reasons for organizations to use AI for personnel selection, has also been addressed in the ongoing discussions among researchers and policymakers. The need for 'Ethical AI,' 'Responsible innovation,' and 'AI alignment,' which include stronger laws and regulations, has been proposed by many scholars (Coeckelbergh, 2021; Pasquale, 2020). Despite the continuous calls for efficient laws and regulations concerning AI, another respondent emphasized that there are no guarantees that AI will be used with great care just because of the creation of new laws, making an analogy to speed limits for cars, which cannot prevent people from breaking laws. And the same applies to AI, 'some people will just not care about the rules.'

4.3. Meta-algorithmic judgments at present (and in the future): Discretion and trust

A recurring theme touched upon during all interviews was related to the importance of human discretion in relation to AI-powered recruitment systems. All Type 1 meta-algorithmic judgments entail a certain level of discretion (more or less regulated by the organization). Here, I will present just a couple of examples to illustrate the fundamental meaning of human trust/distrust in algorithms. One top HR manager working at a Swedish municipality, who had been responsible for implementing an interview robot (*not* machine-learning) at her department, provided an illuminating story about the trust and faith she had in this technology:

> "The thing with this AI robot is that it's completely unbiased. It can observe you and asks you competency-based questions, but it will only evaluate how well your personality traits and performance matches the job profile. There is no arbitrariness whatsoever, because the robot applies the same standards and criteria to all candidates […] Now, one day one of my recruiters told me that she couldn't understand why the robot had given such high scores to one of the candidates, …he had a weak CV and his Swedish was rather poor. She nevertheless decided to follow the robot's advice and invited the guy. I told her that one must trust the robot, and that we shouldn't read CVs before the screening stage. Yeah, and you know what, well…she interviewed the candidate, and he was a fantastic person. We later hired him. The robot didn't have any problem understanding the candidate, despite his foreign accent, and most importantly: the robot doesn't discriminate against anybody. As a recruiter, you have to take advantage of this valuable technology and use your own professional discretion to make the best possible decisions."

This example demonstrates how a meta-algorithmic judgment of the discretionary Type 1 may be divided into compliancy/non-compliancy approaches, depending on how the human recruiter decides to handle the recommendations created by the AI system. The HR manager in this case expressed her great trust in (perhaps even overreliance on) the recruitment robot, while her colleague was more skeptical, although she followed its advice. Trust was the real crux of the matter, but it nevertheless resulted in compliancy. But if

we allow ourselves to speculate a bit, then we could easily imagine quite a different outcome. Suppose the recruiter above had simply refused to follow the robot and decided instead to invite a different candidate (for such scenarios, see Bursell & Roumbanis, 2024). Most organizations expect their recruiters to use their professional judgment and common sense to scrutinize the algorithmic output, thus making their discretion an integrated part of the screening process. Some of the respondents revealed that they sometimes felt rather perplexed about how the AI chatbots (*machine-learning*) make their assessments of the candidates during the screening interviews. For example, one respondent said that she had read through the communication between some of the job candidates and the chatbot and found that 'It clearly had misunderstood what some of the candidates had answered, which then formed the basis for its assessments.' But how much time and effort should a human recruiter really have to spend on controlling the AI advice?

One respondent working for a large recruitment agency took quite a different approach when handling the company's AI recruitment system. The organization had acquired a state-of-the-art chatbot technology (*machine-learning*) to be used for candidate screening. A rather different aspect of meta-algorithmic judgment Type 1 was disclosed in this case, an aspect that clearly highlights a recruiter's distrust in algorithms.

> "The people that provided this chatbot strongly recommend that we make sure that the candidates are always screened by it first before anything else. But I must admit that I prefer to look at the candidates' CVs first, and then send them an invitation with a link to the chatbot. I know this is not the way the providers want their customers to use their system, because you're not supposed to reverse the order like I do (laughs). I guess I have a trust issue when it comes to algorithms."

This statement shows a form of non-compliancy with regard to how this chatbot is supposed to be used to maintain a proper screening routine, thus reflecting what previous scholars have identified as 'algorithm aversion' (Burton et al., 2020; Dietvorst et al., 2015; Prahl & Van Swol, 2017). Still, other respondents answered that they had learned to appreciate the chatbots they worked with, while some mentioned that they often had to remind themselves to not take the scores for granted. There is always a potential risk associated with overreliance in AI recommendations (i.e., 'algorithm appreciation'), because following them 'blindly' may inadvertently lead human decisionmakers astray (Parasuraman & Manzey, 2010; Wagner et al., 2018).

4.4. Imagined AI futures and meta-algorithmic judgments

People often have very different, and sometimes contradictory, expectations about the future, especially regarding new emerging technologies (Konrad, 2006; van Lente, 2012). The expert recruiters I interviewed had quite similar expectations about the future impact of AI on the hiring profession. Basically, most of them thought the assessment process would eventually become much more automatized, and they had all witnessed the increasing role of meta-algorithmic *discretionary* Type 1 judgments in the recruitment profession as a whole. The task of making sense of algorithmic recommendations was commonly understood as belonging to the present. One 55-year-old male respondent working at a large professional hiring company imagined the future of AI in the recruitment industries in the following manner:

> "Well, I think that even twenty years from now, the final candidates will probably still be interviewed by another human. Of course, it's very difficult to speculate, but …I mean, how long will it take to refine the precision of the technologies, and for how long will hiring be based on the symbiosis between humans and AI? On the other hand, I can also imagine that, in the long run, AI will become much smarter in attracting and finding the right candidates independently via social media networks. And to be honest, I don't see any problems in, …well let's say: if a client needs five controllers next week, then an AI will do the work and send five controllers to the client's inbox. In my view, that's absolutely not an impossible scenario. It might sound scary and very futuristic, a bit like George Orwell's *1984*. But I think we're moving in that direction."

For most of the respondents, however, human involvement was still viewed as essential to the hiring process, at least if they looked ten or twenty years into the future, suggesting that we could expect more meta-algorithmic *discretionary* Type 1 judgments in the years to come. The impact of human discretion in relation to AI could, of course, come to change quite dramatically in the long-term future, if new 'hybrid intelligences' (Jarrahi et al., 2022; see, also, Murray et al., 2021) were developed based on mutual augmented human-AI learning processes. The last example, suggested by the male respondent quoted above, can be interpreted as belonging to the *disengaged* Type 3. One 40-year-old female chief recruiter at a mid-sized recruitment company offered an analogy to the aviation industries by envisioning the human recruiters as a 'digital air traffic controllers' who monitor all the automatized recruitment processes in a transparent and safe manner. In her view, the human recruiters would simply follow all the candidates in and out through the AI system, to see how they match against certain parameters and data that are used. This imagined future suggests that we should simply expect to see *monitoring* Type 2 meta-algorithmic judgments in the future, depending on when this kind of AI would be available on the market. Similar views were also expressed by a handful of other respondents, who envisaged the recruiter more as a control agent than a real decisionmaker.

One 55-year-old female recruiter who had been working with HR and hiring for nearly two decades, both in private companies and in the public sector, said the following:

> "I'm convinced that the technologies we're using today will be more refined in the next few years, and we must therefore start to re-think what kinds of competencies will be needed in our organizations. We will probably have to employ software engineers who can have a constructive dialogue with the HR specialists, …I think that, in the near future, relatively speaking, we will just have to let the algorithms find a number of suitable candidates, and I will not have to care about how exactly they found them, I mean, if it's through our internal databases, social media, or LinkedIn, etc., …the *algorithms will just know*. Then, of course, I'll be

> involved in refining and improving the system, and I'll audit the results. But the point is, that it will be entirely automatized before the final candidates reach my desk, and then I'll do deep interviews with these candidates, and they'll feel recognized by a real human being."

This could also be interpreted as an expression of a Type 1 meta-algorithmic judgment, but with a higher degree of automation involved (and human trust). The notion that there were positive aspects of having an even more sophisticated AI in the future was a recurrent view among most of the respondents, even if at least half of them mentioned feeling a bit scared at the same time. But one respondent stood out in the way she presented her imagined future. This 50-year-old female recruiter worked for a big Swedish recruitment company, and she said:

> "I think that AI will take over the entire recruitment process, depending on how far into the future we're looking. We're already using tests to measure personality, IQ, EQ, knowledge, etc., and this can already be integrated into software systems. And if we look even further into the future, then I think you'll just leave some kind of sample or fingerprint, and then it will be scanned by an AI that makes an assessment to see if you're suitable for the job or not. And that's all. …That's how I think it's going to be in the future, it's not what I hope for, but I believe it'll be like that."

She did not say exactly when she expected this to happen, in 15 or 30 years from now, but perhaps that is less important given that nobody can predict the future development of AI anyway. However, her expectation suggests that *disengaged* Type 3 meta-algorithmic judgments could become a reality in the future. But she was not finished yet, as her imagined future could take another twist. She continued:

> "Let's think about this Covid-19 pandemic, for example. It's probably a rather small pandemic in comparison to what might come in the future, and still, it made us start working from our homes. So, if we fast-forward a bit, let's say 50 years from now, then we might have had several global pandemics of much greater magnitudes, and people can't meet in person in the same way anymore. And when they must meet, they will have to step into an artificial world, …and AI systems will be checking before we enter different digital rooms if our personalities match and if we can work together. […] I really don't think this scenario is impossible, in fact, I think it's closer to our own time than many people seem to believe today."

Similar ideas about humans potentially being completely replaced by AI and disengaged from the entire hiring processes were also expressed by a few other respondents, although these ideas were depicted using less dystopic imaginary. At present, the primary motives for pushing AI transformation of the hiring process ahead are the general demands to make it more unbiased, fair, and efficient. That is already happening with algorithm-based recruitment systems that provide expert recruiters with recommendations. As a consequence, for many recruiters, this means that they are currently having to rely on algorithmic assessments before taking the next step in the recruitment process, thus they have to make *discretionary* Type 1 meta-algorithmic judgments. But the question is to what extent the new generations of self-programing AI systems will take over the entire evaluation and decision-making process in the future. What is the likelihood of personnel selection being completely automated? Perhaps environmental problems and/or global pandemics (sooner or later) will force humanity to increase its use of AI anyway, despite the risk of this leading to a total loss of control and de-humanizing consequences.

5. Concluding discussion

For all organizations, decision-making is pivotal in determining memberships, rules, control, and more. If AI eventually decides which individuals should be given membership, the rules to be followed, or how scarce resources should be allocated, then that will most likely change the very nature of social life as we know it. That ever happening, however, would certainly be dependent on political decisions, and the future progress of computer science and engineering, including proper development of 'AI alignment,' that is, building safe AI systems by steering them toward ethically approved goals, value regimes, and principles to benefit the original intentions of their human creators. Promoting fairness, explainability and the prevention of harm are some of the criteria that have often been mentioned in this context (Coeckelbergh, 2021; Pasquale, 2020). But how can computer engineers and software companies ensure that their AI systems will always be trustworthy, once they have been unleashed onto the public to carry out important assessments and decision-making processes (Yampolskiy, 2024)?

The two questions I presented in the introduction were: *What motivates the use of AI in personnel selection today, and how will more advanced AI potentially change the role of human judgment in future hiring*? The picture I got from interviewing my 24 respondents was overall quite an enthusiastic and hopeful one, although some did express more ambivalent feelings and fears. Dystopic views were less prominent, however. Most respondents felt confident that human recruiters would never be fully replaced by AI, although I did find a few exceptions. According to the most common view, which almost every single respondent indicated also applies to their colleagues and the HR community more generally, even if AI were to become extremely sophisticated and sensible, humans could still never afford to lose control over the assessment technologies. Regarding the question of what it is that motivates and justifies the implementation of AI in personnel selection, all respondents mentioned the problem of human bias and the lack of efficiency in traditional hiring. However, the question concerning 'algorithmic bias' influencing recruitment decisions was also highlighted by several respondents as a potential risk that should never be ignored.

The exponential increase in meta-algorithmic judgments in organizational decision-making processes is evident across contexts (Bursell & Roumbanis, 2024). Expert recruiters are one of the groups of professionals who use algorithm-based assessment systems for augmented decision-making. Today, most recruiters who have access to different kinds of AI tools are often allowed to use *discretionary*

Type 1 meta-algorithmic judgments. This means they will form their professional judgments by considering the recommendations provided by the algorithms before taking the next step in the hiring process. Type 1 judgments can be in compliance and non-compliance with the algorithmic system, individual trust being a decisive issue across contexts. The issues of discretion and trust can be related to previous research on 'algorithm appreciation' (Logg et al., 2019; Wagner et al., 2018) and 'algorithm aversion' (Burton et al., 2020; Dietvorst et al., 2015; Prahl & Van Swol, 2017). The final decision-making for Type 1 is always in the hands of the recruiters (or the client). The expert recruiters I interviewed for the present study had used either rule-based AI systems and/or interview chatbots/robots for candidate screening. None of them had worked with video hiring software, yet most were aware that it is being used by HR professionals in other countries.

If we look into the future, the respondents all expected more advanced and refined AI systems to be available, which would lead to higher levels of automation. Here, the *monitoring* Type 2 approaches could be assumed to be the most prevalent decision-making synthesis between human judgment and AI, yet the first part of the assessment processes would simply be taken for granted. Type 2 implies that human judgment would be relegated to the role of a control agent responsible for following up and routinely scrutinizing the flows of candidates in the AI recruitment system. Type 2 presupposes that AI would be allowed to autonomously manage many of the key decision-making procedures. And finally, *the disengaged* Type 3 judgment was perhaps not the type of meta-algorithmic judgment most respondents expected to become a reality, even though a few expressed imagined futures in which human judgment had become totally superfluous in the hiring process. The most likely development, if we listen to the expert recruiters, would probably be some kind of sophisticated AI technologies that elicit Type 1 and Type 2 meta-algorithmic judgments.

The expectation dynamics associated with all kinds of emerging technologies are always dependent on the 'situatedness' of the different actors involved, which can generate interpretative flexibility, not least among the intended users (Konrad, 2006). I have tried to demonstrate in as balanced a manner as possible how the expert recruiters are making sense of AI in personnel selection, viewing their more speculative statements about the future as part of the expectation dynamics flourishing in the HR community today. Both the empirical and the theoretical results of the study can be seen as part of the growing knowledge base concerning the sociology of algorithms and automation, and the concept of meta-algorithmic judgment in particular provides a new theoretical lens for use in future research on human-AI relations (e.g., Hunt & O'Reilly, 2022; Kares et al., 2023; Strohmeier, 2022; van den Broek et al., 2021). The study is based on an alternative theoretical approach that can be contrasted with, for example, the actor-network perspective (Latour, 2005) by assuming a distinct ontological difference between living human cognition and non-living AI.[2] However, this distinction does not in any way disregard the importance of 'common learning processes' and the creation of 'conjoined agencies' of human and AI in everyday life (e.g., Jarrahi et al., 2022; Murray et al., 2021; van den Broek et al., 2021). But meta-algorithmic judgment implies an essential gap between human intelligence and the new types of 'understanding' that can emerge in AI models (Arkoudas, 2023; Bender et al., 2021; Fazi, 2021; Mitchell & Krakauer, 2023).

As we approach the end of the present article, I would like to pose a final theoretical question to conclude this work. What does it mean to say that the very essence of technics is itself nothing technical? Is it a riddle, a philosophical joke, or just pure nonsense? I believe the most charitable interpretation of this seemingly contradictory claim is that it alludes to a deeper existential motive, which is that humans – the very origin of all *techne* – can never be fully absorbed and determined by their own creations. There will always be room for resistance, contradictions, and radical negations at the heart of concrete everyday interactions. On the other hand, what we are witnessing today is how generative AI and deep learning algorithms are transforming some of the most fundamental practical conditions of social life. In fact, these and other emerging technologies are changing not only the *form*, but also in crucial respects the *meaning* of being human today. Hence, there is some truth to the view that the history of technics reflects 'the invention of the human.' And yet, world history has repeatedly demonstrated how human actions and artefacts have the potential to become something essentially 'more-than-human,' which can generate uncertainty regarding the original sources and intentions underlying some life-changing decisions. We must therefore keep in mind what Sophocles proclaimed in *Antigone* (365−366) regarding humanity having 'this resourceful quality of inventiveness – that defines τέχνη – as something clever beyond expectations.'[3]

CRediT authorship contribution statement

Lambros Roumbanis: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of Competing Interest

I have no conflict of interests.

[2] Indeed, the latest efforts in biocomputing to create so-called 'organoid intelligence' by virtue of a synthesis between human stem-cell-derived brain organoids and silicon-based machine learning models could perhaps result in a new and radically different form of 'hybrid intelligence' (e.g., Smirnova et al., 2023; Roumbanis, *forthcoming*). Yet the question of whether these experiments on 'organoid intelligence' will ever lead to a successful merging of the distinct powers of the human brain and generative AI models is impossible to predict; we can only speculate. There are also strong ethical concerns tied to this kind of research that should never be underestimated, especially if we consider the potential suffering that a future sentient 'organoid intelligence' could experience in this world.

[3] The quote is taken from Griffith's translation of Sophocles Antigone (1999). The original Greek text reads as follows: σοφόν τι τὸ μηχανόεν τέχνας ὑπὲρ ἐλπίδ' ἔχων. Griffith does not translate the word 'τέχνη' (or 'τέχνας'), that is, téchne into English.

Acknowledgements

This study was supported by the Swedish Research Council, Sweden, Grant No. 2019-02550. Swedish Research Council for Health, Working life and Welfare, Sweden, Grant No. 2022-00625.

References

Ajunwa, I. (2022). Automated video interviewing as the new phrenology. *Berkely Technology Law Review, 36*, 1173.
Aristotle (350BC/1955) *On Sense and the Sensible*, Oxford, trans. by J.I. Beare.
Arkoudas, K. (2023). ChatGPT is no Stochastic Parrot. But it also claims that 1 is greater than 1. *Philosophy Technology, 36*, 54.
Bareis, J., & Katzenbach, C. (2022). Talking AI into being: the narratives and imaginaries of national AI strategies and their performative politics. *Science, Technology, Human Values, 47*(5), 855–881.
Beckert, J. (2016). *Imagined Futures: Fictional Expectations and Capitalist Dynamics*. Cambridge, MA: Harvard University Press.
Bender, E.M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021) On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? □. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. (pp. 610–623). FAccT '21. New York: Association for Computing Machinery.
Boyatzis, R. E. (1998). *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks: SAGE Publications.
Brayne, S. (2020). *Predict and surveil: Data, discretion, and the future of policing*. Oxford University Press.
Buyl, M., Cociancig, C., Frattone, C., & Roekens, N. (2022) Tackling algorithmic disability discrimination in the hiring process: An ethical, legal and technical analysis. *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 1071-1082.
Burrell, J., & Fourcade, M. (2021). The Society of Algorithms. *Annual Review of Sociology, 47*(1), 213–247.
Bursell, M. (2024). The scope and limits of implicit bias training: An experimental study with Swedish social workers. *Social Science Information, 63*(1), 92–113.
Bursell, M., & Roumbanis, L. (2024). After the algorithms: A study of meta-algorithmic judgment and diversity in the hiring process at a large multisite company. *Big Data Society, 11*(1), 1–18.
Burton, J. W., Stein, M. K., & Blegind Jensen, T. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making, 33*(2), 220–239.
Byrne, D. (2022). A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality Quantity, 56*(3), 1391–1412.
Cantwell Smith, B. (2019). *The Promise of Artificial Intelligence: Reckoning and Judgment*. Cambridge, MA: The MIT Press.
Christou, P. A. (2023). How to use thematic analysis in qualitative research. *Journal of Qualitative Research in Tourism, 3*(2), 79–95.
Coeckelbergh, M. (2021). AI for climate: Freedom, justice, and other ethical and political challenges. *AI and Ethics, 1*(1), 67–72.
Deterding, N. M., & Waters, M. C. (2021). Flexible Coding of In-depth Interviews: A Twenty-first-century approach. *Sociological Methods Research, 50*(2), 708–739.
Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126.
Drage, E., & Mackereth, K. (2022). Does AI debias recruitment? Race, gender, and AI's 'eradiction of difference'. *Philosophy Technology, 35*, Article e89.
Fazi, M. B. (2021). Beyond human: Deep learning, explainability and representation. *Theory, Culture Society, 38*(7-8), 55–77.
Greenwal, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4–27.
Greenwald, A. G., Dasgupta, N., Dovidio, J. F., et al. (2022). Implicit-Bias remedies: Treating discriminatory bias as a public-health problem. *Psychological Science in the Public Interest, 23*(1), 7–40.
Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science Medicine, 292*, Article 114523.
Hoff, J.-L. (2023). Unavoidable futures? How governments articulate sociotechnical imaginaries of AI and healthcare services. *Futures, 148*, Article 103131.
Hou, H. (2021). On the three constraints of the development of artificial intelligence: Value, liberation and responsibility. *Cultures of Science, 4*(3), 159–168.
Howard, A., & Borenstein, J. (2018). The ugly truth about ourselves and our robot creations: The problem of bias and social inequity. *Science and Engineering Ethics, 24*, 1521–1536.
Hunt, W. & O'Reilly, J. (2022) Rapid Recruitment in Retail: Leveraging AI in the hiring of hourly paid frontline associates during the Covid-19 Pandemic. *Digit Working Papers* No.3, University of Sussex.
Iganski, P., & Mason, D. (2018). *Ethnicity, Equality of Opportunity and the British National Health Service*. London: Routledge.
Jarrahi, M. H., Lutz, C., & Newland, G. (2022). Artificial intelligence, human intelligence and hybrid intelligence based on mutual augmentation. *Big Data Society, 9*(2), 1–6.
Jaton, F. (2021). Assessing biases, relaxing moralism: On ground-truthing practices in machine learning design and application. *Big Data Society*, 1–15.
Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A Flaw in Human Judgment*. New York: Little Brown Spark & Co.
Kares, F., König, C. J., Bergs, R., et al. (2023). Trust in hybrid human-automated decision-support. *International Journal of Selection and Assessment, 31*, 388–402.
Konrad, K. (2006). The social dynamics of expectations: The interaction of collective and actor-specific expectations on electronic commerce and interactive television. *Technology Analysis Strategic Management, 18*(3-4), 429–444.
Köchling, A., & Wehner, M. C. (2020). Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research, 13*, 795–848.
Kiviat, B. (2023). The moral affordances of constructing people as cases: How algorithms and the data they depend on obscure narrative and noncomparative justice. *Sociological Theory, 41*(3), 175–200.
Latour, B. (2005). *Reassembling the Social*. Oxford: Oxford University Press.
Lauer, J. (2017). End of judgment: Consumer credit scoring and managerial resistance to the black boxing of creditworthiness. In Raff, & Scranton (Eds.), *The emergence of routines: Entrepreneurship, organization, and business history*. Oxford: Oxford University Press.
Lin, Y. T., Hung, T. W., & Huang, L. (2021). Engineering Equity: How AI can help reduce the harm of implicit bias. *Philosophy Technology, 34*(.1), 65–90.
Lizardo, O. (2022). What is implicit culture? *Journal for the Theory of Social Behaviour, 52*(3), 412–437.
Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes, 151*, 90–103.
Lowry, S., & MacPherson, G. (1988). A blot on the profession. *British Medical Journal, 296*(6623), 657–658.
Mitchell, M., & Krakauer, D. C. (2023). The debate over understanding in AI's large language models. *PNAS, 120*(30), Article e2215907120.
Murray, A., Rhymer, J. E. N., & Sirmon, D. G. (2021). Humans and technology: Forms of conjoined agency in organizations. *Academy of Management Review, 46*(3), 552–571.
Olsson, F. (2024). Implicit Terror: A natural experiment on how terror attacks affect implicit bias. *Sociological Science, 11*, 379–412.
Oomen, J., Hoffman, J., & Hajer, M. A. (2022). Techniques of futuring: On how imagined futures become socially performative. *European Journal for Social Theory, 25*(2), 252–270.
Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors, 52*(3), 381–410.
Pasquale, F. (2020). *New Laws of Robotics: Defending Human Expertise in the Age of AI*. Cambridge, Massachusetts: Harvard University Press.
Pedulla, D. S. (2020). *Making the cut: Hiring decisions, bias, and the consequences of nonstandard, mismatched, and precarious employment*. Princeton: Princeton University Press.
Prahl, A., & Van Swol, L. (2017). Understanding algorithmic aversion: When is advice from automation discounted? *Journal of Forecasting, 36*(6), 691–702.
Raghavan, M., Barocas, S., Kleinberg, J., et al., (2019) Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices. *arXiv* arXiv:1906.09208.

Rivera, L. A. (2015). Go with Your Gut: Emotions and Evaluation in Job Interviews. *American Journal of Sociology, 120*(5), 1339–1389.
Rogers, A., Kovaleva, O., & Rumshisky, A. (2020). A primer in BERTology: What we know about how BERT works. *Transactions of the Association for Computational Linguistics, 8*, 842–866.
Roumbanis, L. (2017). Academic judgments under uncertainty: A study of collective anchoring effects in Swedish Research Council panel groups. *Social Studies of Science, 47*(1), 95–116.
Schiølin, K. (2019). Revolutionary dreams: Future essentialism and the sociotechnical imaginary of the fourth industrial revolution in Denmark. *Social Studies of Science, 50*(4), 542–566.
Shepherd, H. (2011). The cultural context of cognition: What the implicit association test tells us about how culture works. *Sociological Forum, 26*(1), 121–143.
Seppälä, P., & Małecka, M. (2024). AI and discriminative decisions in recruitment: Challenging the core assumptions. *Big Data Society, 11*(1), 1–12.
Smirnova, L., Caffo, B. S., Gracias, D. H., et al. (2023). Organoid intelligence (OI): the new frontier in biocomputing and intelligence-in-a-dish. *Frontiers in Science*. https://doi.org/10.3389/fsci.2023.1017235
Sophocles. (1999). In M. Griffith (Ed.), *Sophocles Antigone*. Cambridge University Press.
Stiegler, B. (1998). *Technics and Time 1: The Fault of Epimetheus*. Stanford University Press.
Strohmeier, S. (Ed.). (2022). *Handbook of Research on Artificial Intelligence in Human Resource Management*. Cheltenham: Edward Elgar.
van den Broek, E., Sergeeva, A., & Huysman Vrije, M. (2021). When the machine meets the expert: An ethnography of developing AI for hiring. *MIS Quarterly, 45*(3), 1557–1580.
van Lente, H. (2012). Navigating foresight in a sea of expectations: Lessons from the sociology of expectations. *Technology Analysis Strategic Management, 24*(8), 769–782.
Wagner, A. R., Borenstein, J., & Howard, A. (2018). Overtrust in the robotic age. *Communication of the ACM, 61*(9), 22–24.
Yampolskiy, R. V. (2024). On monitorability of AI. *AI and Ethics 1-19*. https://doi.org/10.1007/s43681-024-00420-x