OECD Social, Employment and Migration Working Papers
No. 284

Artificial intelligence and labour market matching

OECD

https://dx.doi.org/10.1787/2b440821-en

Organisation for Economic Co-operation and Development

DELSA/ELSA/WD/SEM(2023)2

Unclassified **English - Or. English**

25 January 2023

DIRECTORATE FOR EMPLOYMENT, LABOUR AND SOCIAL AFFAIRS
EMPLOYMENT, LABOUR AND SOCIAL AFFAIRS COMMITTEE

Artificial Intelligence and Labour Market Matching

OECD SOCIAL, EMPLOYMENT AND MIGRATION WORKING PAPERS No. 283

JEL classification : J01, J20, J60, J70

Authorised for publication by Stefano Scarpetta, Director, Directorate for Employment, Labour and Social Affairs.

All Social, Employment and Migration Working Papers are now available through the OECD website at www.oecd.org/els/workingpapers

Stijn Broecke, Stijn.Broecke@oecd.org

JT03511204

OECD Social, Employment and Migration Working Papers

www.oecd.org/els/workingpapers

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcomed, and may be sent to els.contact@oecd.org.

This series is designed to make available to a wider readership selected labour market, social policy and migration studies prepared for use within the OECD. Authorship is usually collective, but principal writers are named. The papers are generally available only in their original language – English or French – with a summary in the other.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Acknowledgments

This publication contributes to the OECD's Artificial Intelligence in Work, Innovation, Productivity and Skills (AI-WIPS) programme, which provides policymakers with new evidence and analysis to keep abreast of the fast-evolving changes in AI capabilities and diffusion and their implications for the world of work. The programme aims to help ensure that adoption of AI in the world of work is effective, beneficial to all, people-centred and accepted by the population at large. AI-WIPS is supported by the German Federal Ministry of Labour and Social Affairs (BMAS) and will complement the work of the German AI Observatory in the Ministry's Policy Lab Digital, Work & Society. For more information, visit https://oecd.ai/work-innovationproductivity-skills and https://denkfabrik-bmas.de/.

Federal Ministry
of Labour and Social Affairs

The report has benefited from contributions from Anna Milanez (on private employment services) and Anne Saint-Martin (on public employment services). Helpful comments and suggestions were received from Angelica Salvi Del Pero, Glenda Quintini, Andrew Green, Marguerita Lane, Chloé Touzet, Anna Milanez, Stewart Butler and Judd Ormsby. This report would not have been possible without the input from the many people listed in Annex A who gratefully lent their time and expertise in speaking to the OECD for this project. Special thanks to the World Employment Confederation (WEC) who arranged interviews within private employment services firms. Members of the WEC also informed the research through a roundtable discussion of how AI is used in private employment services held on 14 October 2021. Thanks also to the participants of the session on "AI for labour market matching" at the International Conference on AI in Work, Innovation, Productivity and Skills held on 21-25 February 2022.

Abstract

While still in its infancy, Artificial Intelligence (AI) is increasingly used in labour market matching, whether by private recruiters, public and private employment services, or online jobs boards and platforms. Applications range from writing job descriptions, applicant sourcing, analysing CVs, chat bots, interview schedulers, shortlisting tools, all the way to facial and voice analysis during interviews. While many tools promise to bring efficiencies and cost savings, they could also improve the quality of matching and jobseeker experience, and even identify and mitigate human bias. There are nonetheless some barriers to a greater adoption of these tools. Some barriers relate to organisation and people readiness, while others reflect concerns about the technology and how it is used, including: robustness, bias, privacy, transparency and explainability. The present paper reviews the literature and some recent policy developments in this field, while bringing new evidence from interviews held with key stakeholders.

Résumé

Bien qu'elle en soit encore à ses débuts, l'intelligence artificielle (IA) est de plus en plus utilisée à des fins d'appariement du marché du travail, que ce soit par les entreprises, les services publics et privés d'emploi, ou les sites et plateformes d'emploi en ligne. Les applications vont de la rédaction de descriptions de poste à l'analyse faciale et vocale lors des entretiens, en passant par la recherche de candidats, l'analyse de CV, les chat bots, les planificateurs d'entretiens et les outils de présélection. Au-delà des possibles gains d'efficacité et économies associés à ces outils, ils pourraient également améliorer la qualité de l'appariement et l'expérience des demandeurs d'emploi, voire identifier et atténuer les préjugés humains. Il existe néanmoins des obstacles à une plus grande adoption de ces outils. Certains obstacles sont liés à la préparation de l'organisation et des personnes, tandis que d'autres reflètent des préoccupations concernant la technologie et la manière dont elle est utilisée, notamment : la robustesse, la partialité, la confidentialité, la transparence et l'explicabilité. Le présent document passe en revue la littérature et certains développements politiques récents dans ce domaine, tout en présentant de nouvelles données provenant d'entretiens avec des parties prenantes clés.

Übersicht

Ob bei privaten Personalvermittlern, öffentlichen und privaten Arbeitsmarkt-dienstleistern, Online-Jobbörsen oder Plattformen – für das Matching von Arbeitsangebot und Arbeitsnachfrage wird zunehmend auf künstliche Intelligenz (KI) zurückgegriffen. Die Anwendungen reichen vom Verfassen von Stellenbeschreibungen über die Rekrutierung von Bewerberinnen, die Auswertung von Lebensläufen, Chatbots, Terminplaner für Vorstellungs-gespräche und Auswahlinstrumente bis hin zu Gesichts- und Stimmanalysen bei Vorstellungsgesprächen. Viele Instrumente versprechen mehr Effizienz und Kosteneinsparungen, könnten aber auch zu einer qualitativen Verbesserung des Matching und der Stellensuche beitragen oder unbewusste Vorurteile aufzeigen und verringern. Es gibt jedoch einige Hindernisse, die einer umfassenderen Einführung dieser Instrumente im Weg stehen. Einige betreffen die Bereitschaft von Unternehmen und Arbeitskräften, solche Instrumente einzusetzen. Andere wiederum rühren von Bedenken in Bezug auf die Technologie und ihre Nutzung her, insbesondere im Hinblick auf Robustheit, Verzerrungen, Datenschutz, Transparenz und Erklärbarkeit. Diese Publikation bietet einen Überblick über die einschlägige Literatur und neue Politikentwicklungen in diesem Bereich und präsentiert neue Befunde aus Interviews mit relevanten Akteurinnen.

Table of contents

Acknowledgments 3

Abstract 4

Résumé 5

Übersicht 6

Executive Summary 9

Principaux résultats 11

Zusammenfassung 13

1 Introduction 15

2 AI Tools for Matching 18
- Optimising job descriptions and CVs 19
- Applicant sourcing and outreach 21
- Job search 21
- Screening and shortlisting 22
- Interviewing 25
- Post-interview negotiations 25
- Help with administrative tasks 25

3 Potential Benefits of Using AI in Matching 27
- Efficiencies and cost savings 27
- Quality of matching 28
- Improving the candidate experience 30
- Addressing human bias 30

4 Adoption Rates and Barriers to Further Adoption 34
- The adoption of AI for matching 34
- Barriers related to organisation and people readiness 35
- Barriers related to concerns about the technology 39

5 Recent Policy Developments with an Impact on the Use of AI in Matching 49
- Transparency and consent 50

Human in the loop and the right to contest 51
Data protection and privacy 54
Bias and discrimination 55
The question of accountability 60

References 61

Annex A. List of Interviewees 79

Annex B. Interview Guides 80

BOXES

Box 1. Data used by AI systems 16
Box 2. AI for Internal Matching 18
Box 3. Automatic Analysis of CVs and vacancies at the French and Flemish Public Employment Services 20
Box 4. *VDAB's Jobbereik* 22
Box 5. Randy: Randstad's intelligent chat bot 23
Box 6. VDAB's Next Step profiling model 24
Box 7. How AI can help alleviate the administrative burden on PES staff 26
Box 8. Pymetrics and the quality of new hires 28
Box 9. Predicting length of unemployment spells using AI at Pôle Emploi (France) 29
Box 10. How Pymetrics addresses bias 31
Box 11. How VDAB, the Flemish Public Employment Service, trains caseworkers and jobseekers in the use of AI 36
Box 12. The evidence is mixed as to whether AI is perceived as more or less fair than humans 47
Box 13. Automation bias and how to address it 52
Box 14. The proposed EU AI Act 58
Box 15. Audits of automated employment decision tools in New York City 59

Executive Summary

Good labour market performance depends in part on the efficiency and the quality of labour market matching—i.e., the process by which workers are matched to jobs. Labour market matching involves a range of steps—from writing job descriptions and vacancies all the way to making offers and salary negotiations, passing through the application, shortlisting and interview stages, amongst others. It covers private recruitment by firms, but it can also refer to the activities of public and private employment services, as well as those of jobs boards and online platforms. It could even allude to matching internal to the firm.

One of the reasons good labour market matching matters, is that it influences the unemployment rate. The harder it is, and the longer it takes, to match workers to vacancies, the higher unemployment will be. Unemployment, in turn, brings costs to individuals (e.g., loss of income, skills depreciation), to firms (e.g., loss of output), as well as to society (e.g., less tax, higher benefit expenditure, lower economic growth). The quality of matches is of equal importance. Skills mismatch can harm productivity and innovation, and result in higher turnover and recruitment costs, as well as lower wages and job satisfaction for the workers involved.

In practice, labour market matching is costly, time-consuming, and suffers from imperfect information as well as bias and discrimination. Improving the efficiency and quality of this process is therefore a key policy priority.

Technology may offer some solutions to improve labour market matching. In particular, a range of Artificial Intelligence (AI) applications have emerged in recent years that can be used at various stages of the matching process. These tools claim they can bring: efficiencies and cost savings, faster and better quality matching using larger applicant pools, as well as ways of improving diversity and addressing human bias and discrimination.

Despite its potential promises, the use of AI in matching still appears to be in its infancy, and there are two sets of barriers that contribute to relatively low adoption rates:

- On the one hand, there are barriers related to the readiness of organisations and people to use such tools, including: management culture and resistance from staff; poor digital infrastructure and data; as well as a lack of skills to work with, or alongside, AI.
- On the other hand, there are barriers related to the technology itself that raise concerns about its use in matching, including: doubts about the robustness of some AI tools; risks to human-centred values and fairness (such as the dehumanisation of the matching process, bias and discrimination, and privacy infringements); and concerns around transparency and explainability.

Policies and regulation can go a long way in fostering the development of trustworthy AI and could also help overcome some of the barriers to its further adoption in labour market matching.

Several relevant policy initiatives are already underway. The EU has proposed the AI Act as well as the directive on working conditions in platform work. In the United States, various states have introduced regulation around the use of AI in hiring. All of these have direct implications for how AI is used in matching. Policy developments so far have focused on:

- Promoting transparency in the use of AI in matching, e.g., by requiring recruiters and other organisations to inform jobseekers and to obtain consent before using AI. One particular challenge in this area is how to achieve consent that is meaningful, given the unbalanced power relationships that exist.
- Ensuring that there is a human in the loop, while avoiding the mere rubberstamping of automated decision-making. The latter can be partly addressed by giving individuals a right to contest automated decisions.
- Guaranteeing privacy, both in terms of the collection of new data by AI tools, as well as protecting individuals from personal information being inferred by AI from social media and other types of big data.
- Fighting bias and discrimination through a range of tools, including: anti-discrimination law; data protection legislation (and, in particular, the right to transparency, to an explanation, and to contest automated decision-making); consumer protection legislation; and the continued monitoring of AI throughout its lifetime (e.g., through audits).

These developments sit within wider efforts to adapt existing, and/or introduce new, legislation to regulate AI more broadly, as well as collective agreements, and attempts at national or international standard-setting and other self-regulatory approaches. These approaches have been discussed extensively in other OECD reports on the ethical risks of using AI in the workplace (Salvi del Pero, Wyckoffi and Vourc'h, 2022[1]) and on AI and social dialogue (Krämer and Cazes, 2022[2]).

Principaux résultats

Une bonne performance du marché du travail dépend en partie de l'efficacité et de la qualité de l'appariement sur le marché du travail, c'est-à-dire du processus par lequel les travailleurs sont appariés aux postes vacants. L'appariement sur le marché du travail implique une série d'étapes allant de la rédaction des descriptions de poste à la formulation d'offres et de négociations salariales, en passant par les étapes de candidature, de présélection et d'entretien, entre autres. L'appariement couvre le recrutement privé par les entreprises, mais peut également faire référence aux activités des services publics et privés d'emploi, ainsi qu'à celles des sites d'emploi et des plateformes en ligne. Le terme pourrait même englober les cas d'appariement interne à l'entreprise.

L'une des raisons pour lesquelles le processus d'appariement sur le marché du travail est important, c'est qu'il influence le taux de chômage. Plus il est difficile et plus il faut de temps pour faire correspondre les travailleurs aux postes vacants, plus le chômage sera élevé. Le chômage, à son tour, entraîne des coûts pour les individus (par exemple, perte de revenus, dépréciation des compétences), pour les entreprises (par exemple, perte de production), ainsi que pour la société (par exemple, moins d'impôts, des dépenses de prestations plus élevées, une croissance économique plus faible). La qualité de l'appariement est tout aussi importante. L'inadéquation des compétences peut nuire à la productivité et à l'innovation, et se traduire par une augmentation de la rotation du personnel et des coûts de recrutement, ainsi que par une baisse des salaires et de la satisfaction professionnelle des travailleurs concernés.

Dans la pratique, l'appariement sur le marché du travail est coûteux, prend du temps et souffre d'informations imparfaites ainsi que de biais et de discrimination. L'amélioration de l'efficacité et de la qualité de ce processus est donc une priorité politique essentielle.

La technologie peut offrir certaines solutions pour améliorer l'appariement sur le marché du travail. En particulier, une série d'applications d'intelligence artificielle (IA) a émergé ces dernières années et peut être utilisée à différentes étapes du processus d'appariement. Ces outils pourraient apporter : des gains d'efficacité et des économies de coûts, une mise en correspondance plus rapide et de meilleure qualité en élargissant les cercles de provenance des candidats, ainsi que des moyens de faire progresser la diversité et de lutter contre les préjugés humains et la discrimination.

Malgré ses promesses potentielles, l'utilisation de l'IA dans l'appariement semble encore en être à ses balbutiements, et il existe deux ensembles d'obstacles qui contribuent à des taux d'adoption relativement faibles :

- D'une part, il existe des obstacles liés à la capacité des organisations et des personnes à utiliser de tels outils, notamment : la culture de gestion et la résistance du personnel ; une infrastructure et des données numériques médiocres ; ainsi qu'un manque de compétences pour travailler avec ou à côté de l'IA.
- D'autre part, il existe des obstacles liés à la technologie elle-même qui soulèvent des inquiétudes quant à son utilisation dans l'appariement, notamment : des doutes quant à la robustesse de certains outils d'IA ; les risques posés aux valeurs centrées sur l'humain et à l'équité (comme la

déshumanisation du processus d'appariement, les préjugés et la discrimination, et les atteintes à la vie privée) ; et les préoccupations concernant la transparence et l'explicabilité.

Les politiques publiques et la réglementation peuvent grandement contribuer au développement d'une IA fiable. Elles pourraient également faciliter à terme l'adoption d'une telle IA dans le domaine de l'appariement sur le marché du travail.

Plusieurs initiatives politiques sont déjà en cours. L'Union Européenne a proposé la loi sur l'IA ainsi que la directive visant à améliorer les conditions de travail dans le cadre du travail via une plateforme. Aux États-Unis, divers États ont introduit une réglementation concernant l'utilisation de l'IA dans le recrutement. Tous ces éléments ont des implications directes sur la manière dont l'IA est utilisée dans l'appariement. Jusqu'à présent, les développements politiques se sont concentrés sur les actions suivantes :

- Promouvoir la transparence de l'utilisation de l'IA à des fins d'appariement, par exemple en exigeant des recruteurs et autres organisations qu'ils informent les demandeurs d'emploi et obtiennent leur consentement avant d'utiliser l'IA. Un défi particulier dans ce domaine est de savoir comment obtenir un consentement significatif, compte tenu des relations de pouvoir déséquilibrées qui existent.
- Veiller à ce qu'il y ait un humain dans la boucle, tout en évitant la simple approbation automatique de la prise de décision automatisée. Ce dernier peut être en partie résolu en donnant aux individus le droit de contester les décisions automatisées.
- Garantir la confidentialité, tant en termes de collecte de nouvelles données par les outils d'IA, que de protection des individus contre les informations personnelles déduites par l'IA des médias sociaux et d'autres types de mégadonnées.
- Lutter contre les préjugés et la discrimination grâce à une gamme d'outils, notamment : la loi anti-discrimination ; la législation sur la protection des données (et, en particulier, le droit à la transparence, à une explication et à contester la prise de décision automatisée) ; législation sur la protection des consommateurs; et la surveillance continue de l'IA tout au long de sa durée de vie (par exemple, par le biais d'audits).

Ces développements s'inscrivent dans le cadre d'efforts plus larges visant à adapter la législation existante et/ou à introduire de nouvelles législations et conventions collectives pour réglementer l'IA plus largement, en parallèle de tentatives de créer des normes nationales ou internationales et d'autres approches d'autorégulation. Ces approches ont été largement discutées dans d'autres rapports de l'OCDE sur les risques éthiques liés à l'utilisation de l'IA sur le lieu de travail (Salvi del Pero, Wyckoffi and Vourc'h, 2022[1]) et sur l'IA et le dialogue social (Krämer and Cazes, 2022[2]).

Zusammenfassung

Gute Arbeitsmarktergebnisse setzen u. a. ein effizientes und gutes Arbeitsmarkt-Matching voraus. Dies ist der Prozess, durch den Arbeitskräfte ihrem Profil entsprechende Stellen finden. Das Matching von Arbeitsangebot und Arbeitsnachfrage erfolgt in mehreren Schritten, die von Stellenbe- und -ausschreibungen über Bewerbungen, Auswahlverfahren und Vorstellungsgespräche bis hin zu Stellenangeboten und Gehaltsverhandlungen reichen. Der Begriff deckt die Einstellungsverfahren von Unternehmen ab, kann sich jedoch auch auf die Maßnahmen von öffentlichen und privaten Arbeitsmarktdienstleistern, Jobbörsen und Online-Plattformen beziehen. Er kann darüber hinaus auch für das unternehmensinterne Matching stehen.

Wichtig ist ein gutes Arbeitsmarkt-Matching u. a., weil es Auswirkungen auf die Arbeitslosenquote hat. Je schwieriger und langwieriger die Abstimmung von Arbeitsangebot und Arbeitsnachfrage ist, desto höher ist die Arbeitslosenquote. Arbeitslosigkeit ist sowohl für die betroffenen Arbeitskräfte mit Kosten verbunden (z. B. Einkommensverlust, Kompetenzentwertung) als auch für die Unternehmen (z. B. Produktionsrückgang) und die Gesellschaft (z. B. weniger Steuereinnahmen, höhere Sozialausgaben, weniger Wirtschaftswachstum). Die Qualität des Matching ist ebenso wichtig wie das Matching an sich. Diskrepanzen zwischen Kompetenzangebot und -nachfrage können die Produktivität und die Innovationstätigkeit beeinträchtigen und eine stärkere Personalfluktuation und höhere Einstellungskosten nach sich ziehen. Für die betroffenen Arbeitskräfte kann ein Kompetenz-Mismatch mit niedrigeren Löhnen und einer geringeren Arbeitszufriedenheit einhergehen.

Das Matching von Arbeitsangebot und Arbeitsnachfrage ist kostspielig und zeitaufwendig und wird durch Informationsdefizite, Vorurteile und Diskriminierung erschwert. Daher ist es dringend erforderlich, die Effizienz und Qualität dieses Prozesses zu verbessern.

Künstliche Intelligenz könnte Lösungen für ein besseres Arbeitsmarkt-Matching bereithalten. In den letzten Jahren wurden mehrere KI-Anwendungen entwickelt, die in verschiedenen Phasen des Matching-Prozesses eingesetzt werden können. Sie versprechen Effizienzsteigerungen, Kosteneinsparungen, ein schnelleres und besseres Matching bei größeren Bewerberpools sowie mehr Diversität bzw. weniger unbewusste Vorurteile und Diskriminierung.

Trotz dieses Potenzials wird beim Matching bislang kaum auf KI zurückgegriffen. Zurückzuführen ist dies auf zwei Arten von Hindernissen:

- Zum einen Hindernisse im Zusammenhang mit der mangelnden Bereitschaft von Organisationen und Menschen, solche Instrumente zu nutzen; Beispiele hierfür sind die Managementkultur und Widerstand vonseiten der Belegschaft, inadäquate digitale Infrastruktur und Daten sowie unzureichende Kompetenzen, mit und neben KI zu arbeiten.
- Zum anderen Hindernisse im Zusammenhang mit Bedenken im Hinblick auf den Einsatz künstlicher Intelligenz beim Matching; zu nennen sind hier insbesondere Zweifel, was die Robustheit einiger KI-Instrumente betrifft; Herausforderungen in Bezug auf menschenzentrierte Werte und Fairness (z. B. Entmenschlichung des Matching-Prozesses, Verzerrung und

Diskriminierung, Verstöße gegen den Datenschutz) sowie Bedenken im Hinblick auf die Transparenz und Erklärbarkeit.

Gesetzliche Richtlinien und Regulierungen können einen wichtigen Beitrag zur Förderung der Entwicklung einer vertrauenswürdigen KI leisten und dazu beitragen, einige der Hindernisse für ihre Anwendung auch im Arbeitsmarkt-Matching zu überwinden.

In diesem Zusammenhang wurden bereits mehrere Maßnahmen ergriffen. So hat beispielsweise die EU einen Entwurf für ein KI-Gesetz und eine Richtlinie zu den Arbeitsbedingungen von Plattformarbeitskräften vorgelegt. In den Vereinigten Staaten wiederum haben mehrere Bundesstaaten Bestimmungen zur Nutzung von KI bei Einstellungsverfahren eingeführt. Diese Bestimmungen haben unmittelbare Auswirkungen darauf, wie KI beim Matching eingesetzt wird. Der Schwerpunkt der Maßnahmen liegt bislang auf folgenden Aspekten:

- Transparente Nutzung von KI beim Matching fördern; beispielsweise indem Personalvermittler und andere Einrichtungen verpflichtet werden, die Bewerber*innen vorab darüber zu informieren und ihre Zustimmung einzuholen. Eine besondere Herausforderung besteht dabei darin, angesichts des bestehenden Machtgefälles eine aussagekräftige Einwilligung zu bekommen.
- Überwachung des Prozesses durch Menschen sicherstellen und das bloße Abnicken automatisierter Entscheidungen vermeiden. Letzteres kann u. a. dadurch verhindert werden, dass den Betroffenen das Recht eingeräumt wird, automatisierte Entscheidungen anzufechten.
- Datenschutz gewährleisten, sowohl im Hinblick auf die Erhebung neuer Daten durch KI-Instrumente, als auch im Hinblick auf die Ableitung personenbezogener Daten durch KI aus sozialen Medien und Big Data anderer Art.
- Vorurteile und Diskriminierung mit einer Reihe von Instrumenten bekämpfen, beispielsweise einem Antidiskriminierungsgesetz, Datenschutzbestimmungen (und insbesondere dem Recht auf Transparenz, Erklärung und Anfechtung automatisierter Entscheidungen), Verbraucherschutzbestimmungen und die laufende Überwachung der KI-Anwendungen während ihrer gesamten Lebensdauer (z. B. durch Audits).

Den Kontext dieser Entwicklungen bilden weitreichendere Bemühungen, bestehende Vorschriften anzupassen und/oder neue einzuführen, um KI umfassender zu regulieren, aber auch Tarifvereinbarungen sowie Bemühungen zur Schaffung nationaler oder internationaler Standards oder andere Selbstregulierungsansätze. Diese Ansätze wurden bereits in anderen OECD-Berichten eingehend erörtert, u. a. in einem Bericht über die ethischen Risiken eines KI-Einsatzes am Arbeitsplatz (Salvi del Pero, Wyckoffi and Vourc'h, 2022[1]) sowie in einem Bericht über KI und den sozialen Dialog (Krämer and Cazes, 2022[2]).

1 Introduction

One of the key ingredients of a well-functioning labour market is the efficiency with which workers are matched to vacancies. When a young person graduates, or someone (in)voluntarily leaves a job: how easily can they find a (new) job that matches their skills, experience and preferences? Similarly, when an employer has a vacancy, how quickly can they find the right profile for that job?

Matching involves a range of steps as well as a number of actors. The process starts with the writing of job descriptions and posting vacancies and ends with the making of an offer and salary negotiations, passing through a range of steps including: applications, screening and interviews. Matching covers private recruitment by firms, but it can also refer to the activities of public and private employment services, as well as those of jobs boards and platforms. In addition, matching can occur within[1] organisations as well as between organisations, and it could even include the acquisition of certain skills to meet the requirements of a specific vacancy.

The efficiency of the matching process matters because it affects the unemployment rate (Blanchard et al., 1989[3]; Mortensen and Pissarides, 1999[4]). The harder it is, and the longer it takes, to match workers to vacancies, the higher unemployment will be. Unemployment, in turn, brings costs to individuals (e.g. loss of income, skills depreciation), to firms (e.g. loss of output), as well as to society (e.g. less tax, more benefits, lower economic growth) (Feldstein, 1978[5]). The quality of matches is of equal importance. Skills mismatch can harm productivity and innovation, and result in higher turnover and recruitment costs, as well as lower wages and job satisfaction for the workers affected (Mcgowan and Andrews, 2015[6]).[2]

In practice, the efficiency and quality of the matching process can be reduced by a number of factors. Workers may lack information about suitable vacancies, partly because it may be very costly and time-consuming for them to find such information. For companies, recruitment is both a lengthy and expensive process, and it is not always easy to identify skills needs or the right candidate because information is limited. Bias on the part of recruiters may further reduce the quality of matches. Public and private employment services also play an important role in labour market matching but, in practice, resources and information are limited, and they will therefore encounter many of the same challenges as individual recruiters.

In the past, human resource managers and employment services have invested in technology, such as computer programmes and algorithms, to automate, accelerate, and improve various stages of the matching process. "People analytics" (i.e., the use of data, statistical and quantitative analysis to drive human resources decisions) have been used for years to help companies improve various aspects of human resources management, including recruitment, and with the goal of achieving efficiencies and cost savings, faster and better matches, a reduction of human bias and error, as well as improvements in the quality of jobs of the workers involved in matching.

---

[1] While many of the tools and the challenges discussed in this paper will apply to within-organisation matching as well, this aspect of matching will receive less attention.

[2] In addition, the more difficult and costly matching is, the less likely employers are to recruit externally, and the lower the quality of matching is likely to be.

AI builds on these previous technological advances and claims to go beyond. Compared to previous technologies, which were largely coded and rigid, and dependent on structured data, AI can be more flexible. AI applications can use large datasets, including unstructured ones, to carry out human-like cognitive tasks (e.g., recognition, event detection, forecasting, …). Moreover, some AI models can learn from data, or even evolve and/or acquire abilities from interacting with data (OECD, 2022[7]). These advances have been possible thanks to: (i) faster computing power; (ii) improvements in algorithms; and (iii) the availability of big data (see Box 1).

Box 1. Data used by AI systems

In recent years, progress in the field of AI and matching has been possible in part to the emergence of new data sources that can be used in addition to more traditional sources of information. Freire and de Castro (2021[8]) distinguish between two types of information: "explicit information" on the one hand – i.e. information consciously entered by the candidate (or recruiter) such as CV-related information – and "implicit information", on the other, which is inferred from the candidate's interaction with the recruiter and the systems used (e.g. click data, questions asked to chat bots, etc.). Freire and de Castro (2021[8]) then identify 4 categories of information:

- *Social media sites* (e.g., Facebook, LinkedIn, …), which cover information taken from: user profiles, connections and interactivity records, and likes and posts.
- *CVs and job posts*, which are the more traditional (and explicit) sources of information used in matching.
- *Behaviour and feedback*, which can include implicit information embedded in: clicks on job posts, time spent reading a certain job description, job posts saved, etc.
- *Other types of information* coming, for example, from questionnaires, geolocation, etc.

According to Freire and de Castro's (2021[8]) estimates, the type of information currently used by e-recruitment recommender systems is: 38% social network; 43% CVs and job posts; 13% behaviour or feedback; and 6% other types).

In addition to the sources of data reviewed by Freire and de Castro (2021[8]), new sources such as those used in image and voice recognition systems are increasingly making inroads into the recruitment sphere.

Source: Freire and de Castro (2021[8])

The platform economy was a front-runner in using AI for matching and demonstrating the opportunities it offers. Ride-hailing companies, for example, have used machine learning algorithms to match thousands of drivers to millions of customers every day. This matching process is entirely data-driven and automated, learning from past experience, to estimate demand at a particular time in a certain location, informing drivers, planning routes, and setting prices (as well as bonuses and incentives) in function of demand and supply.

The use of AI in the platform economy has not always been without controversy, and similar concerns are now being voiced about the wider use of AI in matching. The algorithms used by platforms have been accused of lacking transparency (Aranguiz, 2021[9]), invading workers' privacy (Bacchi and Asher-Schapiro, 2020[10]), and resulting in biased decisions (Kerr, 2020[11]). Similarly, there are risks involved with the increased use of AI by human resource departments as well as by employment services. AI can be a black box and, without knowing how recommendations and decisions are arrived at, it is dangerous to take them at face value—especially in the employment sphere where the consequences for individuals can be significant.

From a policy perspective, concerns about AI need to be addressed since they might affect the efficiency and quality of matching, and they could reduce investment in and the adoption of AI tools. If AI used in matching is not transparent, if it introduces bias into the selection process, and/or invades the privacy of individuals, then the quality and efficiency of matching will be compromised. There will also be reluctance on the part of jobseekers to use these tools, or to be assessed by them. Staff involved in matching might refuse to work with them or to consider recommendations made by AI tools. Similarly, if the regulation framing the use of AI tools is unclear, organisations might shy away from investing in them for fear of litigation, and the possible financial and reputational consequences in case something goes wrong.

The goal of the present paper is to survey the current state of AI in matching, with a particular focus on the barriers to adoption and the risks associated with the use of AI. The paper also reviews the actions taken so far by policy makers to regulate the development and use of AI in matching. Section 2 of the report offers an overview of the uses of AI in labour market matching and Section 3 discusses the benefits. Section 4 looks at the current adoption of AI in matching and analyses the barriers to greater adoption of these tools. The final section of the report (Section 5) offers a brief review of policy action taken so far by countries.

In terms of methodology, the paper combines literature review and information gathered from semi-structured interviews with key stakeholders. More than 200 reports, articles and other sources were consulted and this information was supplemented new insights gained from semi-structured interviews with 23 key stakeholders from public and private employment services, online jobs boards, and human resource leadership and talent acquisition professionals. Annex A provides further detail on these interviewees and Annex B provides the set of questions that were used as a basis for the semi-structured interviews. The objective of the interviews was to test the emerging findings of the literature review and to gain new insights from people at the forefront of developments in the field. Achieving representativeness was not a realistic goal given the qualitative approach taken, but rather gaining a wide range of views spanning a variety of stakeholders involved in matching. The goal was to have at least one representative from each type of stakeholder. Potential interviewees were contacted via LinkedIn as well as with the help of the World Employment Confederation (WEC) in the case of private employment services (PrES). Interviews were held online and lasted approximately one hour.

2 AI Tools for Matching

AI is being developed and used for various stages of the matching process—from the optimisation of job descriptions, to applicant sourcing and screening, all the way to interviewing candidates and making job offers. This section provides an overview of the kind of AI applications[3] used in labour market matching, with examples from the human resources field, private (PrES) and public employment services (PES), jobs boards,[4] as well as the platform economy. The focus is primarily on external matching, although AI can also be used for matching internal to organisations (see Box 2).

Box 2. AI for Internal Matching

The discussion in this paper is focused primarily on the external matching of individuals to jobs – i.e. companies recruiting new staff, employment services matching jobseekers to vacancies, etc. However, the movement of workers within organisations, from one job to another (either through lateral moves or through promotions), is also an important part of labour market matching. Such internal mobility can help fill skill gaps in the organisation, while satisfying employees' desire for career progression and new challenges.

Here, also, AI solutions are emerging. For example, the Blue Matching technology developed by IBM uses natural language and machine learning processes to recommend internal job opportunities to employees, tailored to their qualifications and aspirations (Stolz, 2018[12]). The tool infers the latter from employees' skills, current job role and pay grade, performance ratings, location preferences, as well as other internal digital footprints, such as professional blogs. Employees can validate any skills and competencies associated with previous jobs, held within or outside the organisation they currently work for.

One benefit of IBM's Blue Matching is how it learns from new matches and feedback from individuals to improve over time. For example, when new career moves or job roles emerge within the organisation,

[3] One important caveat to this discussion (and the rest of the paper), is that it is not always easy to distinguish "real" AI from more traditional algorithms and software—especially since it is an attractive marketing strategy for developers to sell products as "AI" even when they are not. One representative from a PrES interviewed as part of this project said that, "One third of developers admit AI is just marketing. It's AI in name, but in reality there is no AI." In theory, AI is different from previous technologies in its ability to adaptively interpret and even learn from large data sets to perform non-routine, human-like cognitive tasks—as opposed to more traditional algorithms that are fixed and pre-programmed, and tend to be able to do routine tasks only. Moreover, there are conflicting definitions of AI that make classifying tools as AI difficult. While every care has been taken to focus on AI solutions in this paper, it is possible the discussion includes some non-AI applications. That being said, many of the benefits and risks of AI also apply to these older technologies (see also Giermindl et al. (2021[75]). In fact, while AI creates new challenges for policy and regulation, it would be a mistake to treat it entirely differently from previous technologies as well as from more traditional recruitment and matching practices (see Section 5).

[4] Firms increasingly rely on social media platforms and digital services, such as Facebook, Instagram, LinkedIn, Xing, Monster, and CareerBuilder, to advertise job vacancies and to find well-fitting candidates (Köchling and Wehner, 2020[167]).

these new trends are immediately integrated into the system. The tool also gets to know individuals better, as it monitors how helpful its advice proves to be. Another benefit of the tool is that it highlights opportunities that workers may not necessarily have considered otherwise (Hardy-Vallée, 2019[13]).

In 2018, about 50 000 IBM workers were using Blue Matching on a voluntary basis and, amongst those who received a new job or a promotion, nearly three in ten were assisted by the system (Rosenbaum, 2019[14]); (Lewis, 2019[15]).

More generally, however, the current use of smart apps to foster internal mobility is still modest: almost two thirds of HR professionals surveyed have not yet adopted such tools while only 6% of them report using AI solutions moderately or to a great extent (Zhang, Feinzig and Hemmingham, 2018[16]). This is despite the fact that such tools might, in a way, be simpler to adopt than tools used in external recruitment. One HR Adviser interviewed as part of this project, and who was sceptical of AI in general, conceded that the use of AI for internal matching might be more reliable because it is used within a certain context.

Optimising job descriptions and CVs

There are a number of tools on the market that help recruiters write and optimise job descriptions. The drafting of job descriptions is often the first step in the matching process and it is an important one because it determines the set of information that is available to candidates and whether or not they choose to apply. AI tools that help with this stage of the matching process tend to use data from large quantities of published vacancies which are constantly being updated. Using techniques such as Natural Language Processing, Deep Learning and Word Embedding, these tools analyse the content of CVs and job vacancies, as well as that of occupational taxonomies, and convert the information into structured data that can be linked together in various ways. Currently, such AI tools promise to help employers with: search engine optimisation; the identification of skills, qualifications and responsibilities typically associated with a certain job title; the readability of individual job advertisements, as well as the uniformity of language and branding used across multiple job advertisements.

Public and private employment services also use AI tools to optimise job descriptions. One PrES firm interviewed as part of this project said they use a third-party job description optimisation tool that ensures a number of criteria are met before a job advertisement is posted. These criteria include: sufficient simplicity of language, clear identification of important elements for the candidate, inclusion of the firm's branding, and language candidates tend to like. A recruiter inputs a draft job description into the tool and is offered suggested revisions, which are either accepted or rejected at the recruiter's discretion. The performance of this tool was measured according to the change in the conversion rate – i.e., the portion of applications received relative to overall job openings – and the change in the number of applications. Over the course of the implementation of the tool, the conversion rate increased 14% and the number of applications increased 18% (Textmetrics, n.d.[17]). In Flanders (Belgium), the Public Employment Service VDAB uses tools that, on the one hand, identify skills and competencies that an employer may consider adding to the job description (Box 3) and, on the other hand, identify the occupation that best matches a job vacancy. More precisely, the latter tool (Occupation-finder) estimates the vacancy's "distance" from each of the 600 occupations included in the VDAB's taxonomy of occupations and competencies in order to (re)label the job ad with the "closest" matching occupation.

Many of these AI tools can help recruiters avoid language which might deter certain candidates (e.g., women, minorities) from applying—and so help to promote greater diversity in applicants. Research has shown that men and women react very differently to words used in job vacancies, without necessarily being conscious of it. For example, when job advertisements are constructed to include more masculine than feminine wording (such as "leader", "competitive" and "dominant"), women find these jobs less attractive (Gaucher, Friesen and Kay, 2011[18]). Avoiding such language can help recruiters reach a more diverse applicant pool. The French PES, Pôle Emploi, uses an AI tool that searches job offers for terms and conditions that could be problematic/illegal, such as age or gender requirements.

There are also AI-powered tools that help candidates build CVs and write cover letters. These tools tend to draw on large databases of CVs and job vacancies to look for similarities in skills and experience, as well as job interests, and then recommend keywords to include that will make a candidate's CV or cover letter more likely to be picked for an interview. AI can also help jobseekers save considerable amounts of time. Instead of completing long, online application forms, natural language processing techniques can parse unstructured documents like CVs, extract the relevant information, and automatically complete the forms. Unilever, for example, asks applicants to submit their LinkedIn profile and the AI system combs through the candidate's profile to complete the application for them (Black and van Esch, 2020[19]). Some AI tools can also improve the quality of a candidate's CV and cover letter, for example by identifying skills that they possess but may have forgotten to include. These tools infer skills based on skill bundles contained in historical data. The separate identification of inferred skills allows a recruiter to follow up with the individual to determine whether those skills are indeed part of their qualification. Both the French PES (Pôle Emploi) and the Flemish one (VDAB) use such tools (Box 3).

Box 3. Automatic Analysis of CVs and vacancies at the French and Flemish Public Employment Services

Automatic CV Analysis at Pôle Emploi

The French PES, Pôle Emploi, uses a tool called "Automatic CV Analysis" (*Analyse Automatique de CV – AACV*) which helps jobseekers from the moment they create a profile with *Pôle Emploi*. Based on the CV uploaded by the jobseeker, AACV suggests skills that the jobseeker is likely to possess but did not mention on their CV. The tool does this based, on the one hand, on an analysis of the individual's job history and, on the other, on the PES taxonomy of occupations and skills (ROME). The case worker can validate or reject the suggestions made by the tool, following an interview with the jobseeker. 68% of those using the tool say they are satisfied with the data retrieved from the CVs, the suggestions made regarding which skills to highlight, and the identification of skills they had not mentioned explicitly on their CV.

VDAB's Competency-Seeker

The VDAB's Competency-Seeker helps both jobseekers and employers enrich and refine the skills profiles they have and are looking for. As with Pôle Emploi's AACV, the VDAB tool does this based on the PES taxonomy of occupations and skills (Competent) and the information provided online by individuals. The tool is also used by case workers when helping vulnerable jobseekers identify the skills and competencies they possess. The Competency-Seeker analyses the content of the CV or job advertisement that users have uploaded on the PES website, and/or the information they have entered manually when completing their online profile. The tool then suggests skills and competencies that were not mentioned explicitly in the uploaded document or entered online. For jobseekers, the tool will detect implicit skills (e.g. a "truck driving licence" from any profile which lists "truck driver" as work experience). For recruiting firms, the Competency-Seeker will suggest a set of skills and competencies that could be needed for the job posted, based on the information provided on the job duties, work tasks, and the firm's activities more broadly.

Applicant sourcing and outreach

AI allows employers who are recruiting to cast a wider net and to "headhunt" candidates best suited to the job. Once an employer has put together a job description, the next step in the matching process consists of reaching out and finding potential candidates. Traditionally, vacancies were posted in newspapers or spread through word-of-mouth. This meant that employers had little control over who saw vacancies, and also that they reached relatively few individuals. Over the years, technology has helped employers cast a wider net through, for example, the internet and online job boards like Monster, Indeed and Reed. AI is helping to further improve these types of tools. Companies like Facebook and Google have developed tools that try and predict who might be interested in a particular job advertisement based on their online behaviour and click history, as well as their career trajectory as recorded on platforms like LinkedIn. This means job adverts are better targeted than they previously were. These approaches also mean that AI can assist employers in reaching out to passive candidates by placing job advertisements on platforms and social media where they will be seen by individuals who were not necessarily looking for a job.[5] In addition, some tools, like Entelo, try to predict how likely an individual is to leave their current job by analysing data on their employer (recent layoffs, mergers, stock market value) so that they can be actively targeted by recruiters (Bogen and Rieke, 2018[20]). Entelo does this by searching (scraping) 200 million candidate profiles from 50 Internet sources (Chamorro-Premuzic et al., 2016[21]). Finally, AI can help outreach to specific target audiences, such as racial minorities or certain professions. For example, one PrES firm interviewed as part of this project used a programmatic advertising tool which autonomously identified relevant sites to post job advertisements to, such as niche job boards or community pages.

Public and private employment services are also using AI-based tools to identify and contact potential candidates for existing vacancies. The Flemish PES, VDAB, uses a tool called Talent API which selects the candidate profiles that are most similar to those the employer has already looked at. Each time an employer arrives on the PES website to consult a candidate resume (for example via a Google search), the application displays similar jobseeker profiles the employer can consult. This enables employers to navigate the VDAB website in an intuitive manner and to progressively refine their search to find the most suitable candidate.

Job search

For individuals, AI tools can help personalise the job search process and result in tailored job recommendations. Algorithms can measure similarities between jobseeker profiles and vacancy descriptions to suggest potential matching opportunities, listed in order relevance (i.e. ranked according the estimated degree of similarity). In addition, they can learn from the click history and other actions of jobseekers to show them vacancies that are better matched to their skills, experience and interests. This is called "content-based filtering". The VDAB application Talent API mentioned above provides jobseekers with a shortlist of job vacancies that best match their profile. Some tools also use "collaborative filtering", which aims to predict the kind of vacancies someone might be interested in, based on what people similar to that person are looking at (Bogen and Rieke, 2018[20]).

Some tools identify skills gaps to help jobseekers improve their employability. Most tools look for matches between job vacancies and existing skills. However, some tools go beyond that and identify skills gaps, as well as potential training, to jobseekers. For example, the Flemish PES, VDAB, uses a tool called *Jobbereik* (Job Reach) which aims to encourage jobseekers to broaden their job search while providing guidance for potential career moves (Box 4). A representative from a PrES interviewed as part of this

[5] One challenge here is individuals who are not currently "online" in any way and cannot, therefore, be reached through social media or other digital means. PES may have better access to such individuals and could also play a role in getting them "online".

project said they used a similar tool: "Instead of the tool simply looking at your existing skills and certifications and finding the job you currently match, it looks at the next step in your career and shows you where you could get to if you acquired certain recommended courses and certifications as well as work experience. You could not do this manually for such a large number of people, but with AI you can, and people can fulfil their potential, get higher salaries, etc." A managing director for workforce transformation at a large consulting company added: "AI can increase transparency. It does not only help individuals identify great matches, but it also helps them understand why there is a match, what skills they have, what skills they lack, and how they can chart a course towards skills development. Indeed, part of the promise of AI is its ability to help individuals build skills and then get a better job."

Box 4. *VDAB's Jobbereik*

VDAB's *Jobbereik* application helps users visualise the set of occupations and jobs that they could consider given their existing skills and competencies. The basic premise of *Jobbereik* is that core skills and competencies are partially transferable across occupations and industries. In practical terms, users enter their desired occupation to receive a list of possible alternative occupations that require more or less similar skills profiles. Given that many people focus only on the occupation last exercised when searching for a new job, *Jobbereik* may open up new employment opportunities by highlighting occupations that they would otherwise not have considered. While some of the suggested occupations will be readily accessible to the jobseeker, others will require the acquisition of new skills. In each case, *Jobbereik* provides insight into the skills gap that needs to be closed in order to make the corresponding career move, drawing on the individual's skills profile. This helps gauge the adequacy and feasibility of the proposed career moves, not only according to the occupational preferences of individual jobseekers but also according to the type of skills they are ready and able to learn. To further support job mobility, VDAB is developing a new functionality that will enable *Jobbereik* to suggest a list of possible education and training programmes for each proposed career move.

Screening and shortlisting

Screening and shortlisting are some of the more time-consuming parts of the recruitment process. Once applications have been received for a specific position, the recruiter may need to do some screening to keep only those applicants who have the essential requirements for the job. This might include: checking CVs against job prerequisites, carrying out background checks, and possibly running some tests/assessments. In addition to (and often in parallel with) basic screening, recruiters will shortlist those applicants who correspond best to the profile sought. Both these processes are very resource-intensive and, with the increased number of applications resulting from digitalised and online recruitment processes, many recruiters struggle to keep up with the workload. As a Workforce Strategy & Transformation Leader at a large consulting company interviewed as part of this project said: "With the volume of applications received, it is challenging to go through each resume."

AI has vastly expanded the range of time- and resource-saving possibilities when screening and shortlisting candidates. For years, technology has helped automate some of the tasks involved in screening and shortlisting—e.g., by searching for keywords in a CV or reviewing the answers to a list of pre-set questions. Building on these technologies, AI has expanded the range of possibilities. In particular, AI is less rigid than previous technologies were. In the past, for example, an exact match would have been required between a word used in a vacancy and that used in a CV. However, through "semantic expansion", AI can parse CVs and take a single word such as "accountant" and expand the information linked to the candidate to include known synonyms, such as "account specialist." This ensures that candidates are not ruled out by narrow phrasing or phrasing that is slightly different from advertised text.

There are many such parsing tools on the market today to help employers review CVs and rank candidates on how well their qualifications, skills and experience match the job description, as well as the employer's past screening and hiring decisions. Chat bots can also be used for screening, such as Mya or Randy (see Box 5), which analyse a candidate's replies and make recommendations about who should move to the next stage of the recruitment process (Bogen and Rieke, 2018[20]). Other AI tools field tests and run interactive games with candidates, which are then analysed automatically to uncover patterns that can identify successful workers based on their cognitive, social and emotional traits (Bodie et al., 2017[22]) (Sánchez-Monedero, Dencik and Edwards, 2020[23]). Finally, there are tools that allow companies to perform background checks on applicants, including by looking at their social media activities (a procedure sometimes called "cybervetting" (Chamorro-Premuzic et al., 2016[21])). Fame, for example, looks for instances of misogyny, bigotry, racism, violence and criminal behaviour in publicly available online content, and flags these to the recruiter—although some social media platforms have now restricted access to their data for the purpose of employment decisions or surveillance (Bogen and Rieke, 2018[20]).[6] An HR Innovation Strategist at a large consulting company interviewed as part of this study said: "In the United States, it is relatively common to scrape social media for background checks. Companies don't want to hire someone who could be an embarrassment."

Box 5. Randy: Randstad's intelligent chat bot

Since 2019, the human resource consulting firm, Randstad, has used an intelligent chat bot, called Randy, to pre-select candidates. Discussions with candidates last the same time as an in-person pre-hiring interview (approximately 20 minutes), during which Randy collects information on technical and behavioural skills required in a certain job opening, without the need for the candidate to submit forms or CVs, or speaking to a human being. Using a learning algorithm, Randy optimises the job search and the match between a profile and a position. Available on Facebook Messenger and on the Randstad website, Randy gamifies the application process through multiple choice questions and games adapted to each business. At the end of the exchange with the job candidate, Randy offers a scoring of the evaluated application to provide Randstad consultants with the profiles that best meet their recruitment needs. If the candidates agree, they are contacted again within 48 hours by a recruitment consultant. This allows consultants to focus on the most strategic dimension of their role: the interview and final evaluation of applications.

Source: https://www.illuin.tech/en/projects/randy-2/

Private employment services also use AI tools to save time and money when screening and shortlisting jobseekers. In discussions with PrES, one firm said it had a chat bot that engages with job candidates after they apply for a job to assess whether they meet certain minimum requirements for the position, such as the possession of a driver's license. Job candidates who do are then prompted to schedule time with a recruiter. The point of the tool is to eliminate the need for recruiters to screen based on minimum requirements themselves, by phone, and to speed up the review process for candidates. In this firm, 20% to 30% of interviews that recruiters have with job candidates are scheduled for the same day using the chat bot. Several PrES interviewed as part of this project said they used recommender tools, i.e., tools that would offer a recruiter a list of top candidates, thereby replacing the need for the recruiter to

[6] The particular example that Bogen and Rieke (2018[20]) refer to is based on a 2018 incident where an automated baby-sitter rating system used Facebook and Twitter data to predict the risk of drug abuse, bullying, or "having a bad attitude". While this company was allowed to access personal data, it violated the platforms' use of personal data to evaluate person for decisions on hiring or eligibility.

conduct manual searches. Similarly, for online job platforms that offer firms the promise of a curated pool of top talent, AI can be used to curate prospective job candidates and inform the decision to admit them to the platform. In the past, this process was manual: hiring managers would review profiles and approve or deny candidates' entry. Now, machine learning AI tools are capable of reviewing profiles and, using historical data on attributes of candidates who have succeeded in the past, assign scores indicating candidates' likelihood of success. Those with sufficient scores are automatically accepted, while those with low scores are flagged for further review by a hiring manager, who makes a final decision.

Public employment services use AI for screening as well as profiling jobseekers. The Flemish PES VDAB aims to reach out to and screen all new jobseekers within six weeks of registration with the PES, while giving priority to those facing the greatest labour market difficulties. As part of this contact strategy, an AI-based profiling model – known as "Next-Step" – assists caseworkers in identifying vulnerable jobseekers (Desiere, Langenbucher and Struyven, 2019[24]). The tool classifies jobseekers into five groups according to their estimated probability of (re)employment (i.e., their profiling scores), so that caseworkers can contact them by order of priority (Box 6).

Box 6. VDAB's Next Step profiling model

VDAB's Next Step is a profiling model that uses AI to identify jobseekers who face the greatest labour market difficulties.

More specifically, Next Step uses a random forest model to estimate the probability that an individual finds a job within the next six months. The data used to estimate these profiling scores are collected from the jobseekers when they register at the PES and complete their online profile. These include three main categories of information:

- Socio-economic characteristics, such as place of residence, age, education level, vocational skills and competencies, as well as previous work experiences and unemployment spells (including the current one);
- Work preferences, for example in terms of occupation, industry, or location;
- Behavioural indicators, using jobseekers' activity on the VDAB website as a proxy for job search behaviour. For example, the following "click data" are collected: logging in, adding or modifying information on the My Career user interface, clicking on job vacancies, etc.

Jobseekers, as well as the caseworker assisting them in their return to work, can complement and update these data at any time throughout the job search period. The database that feeds the prediction model is updated on a daily basis, and so are the profiling scores.

The only information that caseworkers receive is the group to which a jobseeker belongs, individual scores are not provided in order not to influence their decision about what to do next. Caseworkers decide what kind of support a jobseeker needs after a phone interview, drawing on their expertise and their own evaluation of the individual's circumstances.

Unemployed persons whose profiling score cannot be estimated are contacted first. For some, this results from specific circumstances under which they can be exempted from the requirement to seek work, for example when they are enrolled in training programmes. In other cases, the jobseeker profile does not contain all the necessary information for the model to run properly, which may indicate that the person needs guidance on how to complete his/her online profile.

Once caseworkers have contacted all individuals belonging to this first group, they can start dealing with the other four groups of jobseekers.

Interviewing

AI solutions are also being introduced to help save time and improve candidate assessment at the interview stage. This is, arguably, one of the stages of the matching process that is the hardest to automate and where human interaction and skills will continue to remain important. There are nonetheless tools that claim they can help recruiters analyse data collected during the interview (e.g., word choice and complexity, tone, eye contact, mood and facial expressions). Interviewer.AI, for example, uses computer vision, natural language processing and audio analysis to assess the pace, body language, dressing, eye contact, facial motion, etc. of candidates to score them on communication, professionalism, sociability, and attitude. Through a combination of industrial organisation psychology heuristics and machine learning, the tool provides an interview score for each candidate. This is additional information that the interviewer can take into account when making a recruitment decision. HireVue is another example of a company that offers tools to carry out automated interviews to evaluate candidates. Note that there are also tools available that can be used by job candidates to help prepare for interviews, e.g., to gain training in transversal skills like public speaking (Verhagen, 2021[25]).

Post-interview negotiations

Some AI tools assist employers in post-interview negotiations. Once a candidate has been picked, there can be further negotiations around salary, benefits etc. before a candidate accepts a job. This is an important part of the matching process, since it will determine whether the match will be sealed or not. If the package is not attractive enough, the candidate may not accept the job offer but, at the same time, the company will not want to offer more than the candidate will have expected. Companies like Beqom claim they can optimise and benchmark compensation based on candidate characteristics; they offer natural language processing to analyse candidate sentiment in written correspondence and adjust compensation offers accordingly. Bogen and Rieke (2018[20]) describe how Oracle's recruitment software adjusts salary, bonus, stock options, and other benefits based on predictions of how likely a candidate is to accept a job offer. All these tools tend to be constantly updated with the latest data, from which they learn, and also produce predictions for future salary developments.

Help with administrative tasks

Throughout the matching process, there are a range of administrative tasks that AI can help automate. These are tasks that often take up of a lot of staff time but do not add much to the quality of the matching itself. Some tools help with the organising and scheduling of interviews. Others help in communication with candidates, automating the answering of basic questions that come back again and again. While a Q&A page on a website can help employers save some time, chat bots can offer a more personalised and flexible solution, particularly when they use natural language processing to interpret questions and can learn from feedback in order to improve their answers over time. Chat bots can also provide immediate feedback to applicants whose skills or qualifications do not meet the job requirements, so they are not left hanging on. Chat bots have the added advantage that they can allow employers to collect information on applicants which (as discussed above in the case of the Mya and Randy chat bots) can help with the selection procedure. Public employment services are also increasingly using smart bots to help case workers handle the many requests for information and assistance they receive every day (Box 7). Some tools are designed to be used by customers in self-service mode, answering basic questions and helping with administrative procedures. Others are for internal use and assist case workers in dealing with specific queries or help them manage their email inboxes. Finally, online jobs platforms and professional social networks like LinkedIn use AI to detect fraud and scams, and to identify redundant or duplicate job openings. These are tasks that would take a considerable amount of human labour but add little to the efficiency or the quality of the matching itself. As Joaquin Quinonero Candela from LinkedIn said, “AI can be leveraged to make hiring safe, trusted, and professional.”

Box 7. How AI can help alleviate the administrative burden on PES staff

As public employment services (PES) provide a wide range of financial support and employment programmes, case workers have to deal with numerous queries from jobseekers and employers regarding their rights and obligations vis-à-vis the PES. AI-based tools can help with these routine and repetitive tasks, freeing up time for case workers to concentrate on their core mission: helping jobseekers to find jobs and employers to fill vacancies.

The added value of these new tools was particularly visible during the COVID-19 crisis, at a time when PES worldwide faced a rapid surge in unemployment. Against this backdrop, the deployment of smart bots took a leap forward in a number of countries and helped PES staff handle a massive influx of questions—in part related to the newly implemented COVID-19 measures (Glassclock, 2020[26]).

In the United States, the Texas Workforce Commission (TWC) developed and rolled out a chat bot shortly after the start of the COVID-19 crisis, which has served over two million people and answered over nine million questions during the pandemic (Accenture, 2021[27]; Center for Digital Government, IBM and NASCIO, 2021[28]). On average, the tool (known as “Larry”) answers requests in just over four messages, but it also offers the opportunity for the user to request to be contacted should he/she need more information.

Smart bots have also been developed for internal use only, to help PES staff answer specific queries or manage their email inboxes more efficiently. In Arizona, the Department of Economic Security uses a chat bot that helps staff navigate the rules and regulations that determine eligibility for unemployment benefits (Center for Digital Government, IBM and NASCIO, 2021[28]). The tool enables call centre employees to better serve claimants as they can find the information they need faster and easier.

In France, Pôle Emploi uses an AI-based tool (called “Contact via Email”) which sorts messages received from customers and recommends pre-defined answers for the case worker to send back. The case worker can then modify or personalise the proposed message, use it as it is, or reject it altogether.

3 Potential Benefits of Using AI in Matching

AI promises many benefits when it comes to the efficiency and the quality of matching. This section discusses some of these potential benefits, which range from efficiencies and cost savings, to faster and better quality matching using larger applicant pools, improving the candidate experience, as well as a possible reduction in human bias. Ultimately, these benefits should improve the matching process in the labour market, resulting in lower unemployment, as well as higher productivity and growth.

Efficiencies and cost savings

Matching jobseekers to vacancies is both time-consuming and costly. According to LinkedIn (2016[29]), over a third of companies claim that a limited budget is a top challenge for recruitment. In the UK, Glassdoor (2020[30]) estimated that the average employer spends about GBP 3 000 and 27.5 days to recruit a new worker. This is very close to an estimate for the United States, with a cost-per-hire of approximately USD 4 000 and a time-to-fill of 42 days on average (SHRM, 2016[31]).[7] Leong (2018[32]) reckons it takes a recruiter an entire day to review 100 CVs, which is an ineffective use of time considering that 75-80% of applications are estimated to lack the right qualifications (Benfield, 2017[33]). Moreover, the demands on recruiters and other staff involved in matching has increased with digitalisation, which has lowered the barriers for individuals to apply for a job. Now, the marginal cost of applying for an additional job is very low, and companies are flooded with applications. The demand on the time of recruiters has therefore increased and a Workforce Strategy & Transformation Leader at a large consulting company interviewed as part of this project indicated that "Recruiter burnout is a very big issue." Similarly, an HR Adviser in a large manufacturing firm who has so far been sceptical of AI said: "At some point, we encounter a dilemma. On the one hand, you worry that AI might not be fair and transparent. On the other hand, if you receive 100 000 applications, you can't realistically give every applicant a fair chance."

AI vendors are keen to advertise the alleged benefits of their solutions, but independent estimates of the effectiveness of these tools do not appear to exist. For example, Skillate argues that its customers have seen a 40% fall in the cost of hiring and a 65% reduction in the time to hire. XOR claims 33% faster recruitment and 50% more efficiency. HireVue states that its software makes hiring three times faster and increases interview show rates by over 20%. These statistics are all part of firms' marketing strategies and should be taken with a grain of salt. As the Manager of Digital Ethics at a large consultancy company explained, "Clients often want to save money, but also to add scale, do things faster, more efficiently, more accurately. AI makes promises in those areas, however we don't look enough at what the real impact of AI is. There are no double-blind experiments, where we look at the outcomes with and without AI." That being said, many HR professionals seem to agree that the use of AI in recruitment is likely to be time-saving (Hekkala and Hekkala, 2021[34]) and job candidates appear to think so too (Kim

[7] Time-to-recruit and time-to-fill are two slightly different concepts, with different staring points. Time-to-fill starts from the decision to open up a new role. Time-to-recruit/hire starts from the time of application.

and Heo, 2021[35]). There is some fear of losing out among companies who do not adopt AI. A managing director for workforce transformation at a large consulting company interviewed as part of this project said: “For many organisations, the greatest risk is that they don’t adopt AI and that they are outpaced by others who are able to get insights into talent to help them be more competitive.”

In the context of employment services, there is some evidence that AI speeds up job search. Korea’s public jobs portal network introduced “The Work”, which uses AI technology to analyse jobseeker data (e.g., CV, training received, areas of interest, location) and to provide employment information that is tailored to each user, based on the 'National Job Information Platform'. Instead of having to carry out individual searches, the jobseeker receives tailored job opening recommendations. While jobseekers had previously spent an average of 10 minutes searching for job-related information on other, separate sites, “The Work” provided the same information within 5 seconds of log-in (OECD, 2018[36]). Among PrES, Glen Cathey from Randstad said, at a session of the OECD 2022 AI-WIPS Conference, that their chat engine led to a 50% decrease in the time it takes to submit talent to hiring managers.

Evidence from the platform economy suggests that algorithms can bring efficiencies. There is some evidence that platforms are more efficient in matching and have created new employment opportunities (Schwellnus et al., 2019[37]). It has been estimated that the enhanced efficiency that platforms bring could result in an additional 72 million jobs worldwide and spur global GDP by 2% within a decade (Manyika et al., 2015[38]).

Quality of matching

There are a number of mechanisms through which AI could improve the quality of matching. First, AI makes external recruitment quicker and cheaper, which provides companies with improved access to external talent which, in turn, could lead to better matches since companies are no longer relying on internal talent only to fill vacancies. Second, AI could increase the quality of matches by improving the data available to decision-makers (see Box 8). AI could act as a check on decisions made by human beings (WEF, 2020[39]), but it could also provide recommendations or advice which humans then act upon. Finally, by freeing up the time of staff involved in matching, they will spend less time on administrative tasks that add little to the quality of matches, and spend more time on higher value-added tasks like rapport-building, interviewing and negotiating.

Box 8. Pymetrics and the quality of new hires

The company Pymetrics has developed a gamified assessment tool which collects cognitive and behavioural data to allow companies to assess the soft skills of job candidates. This tool has been inspired by cognitive and behavioural science, which has established a link between people’s aptitudes, characteristics and personality traits, on the one hand, and job performance, on the other. Pymetrics assessments provide employers with additional information over and above that contained in more traditional sources, such as an applicant’s CV.

The Pymetrics assessment consists of 25 minutes of behavioural exercises that assess candidates’ decision-making, generosity, learning, quantitative reasoning, effort, fairness, attention, numerical agility, focus, risk tolerance, and emotion. Prior to taking the test, candidates are informed what they will be assessed on. The results are used to provide employers with a three-tiered recommendation: candidates are either highly recommended, recommended, or not recommended. The individual scores are not communicated to the employer.

A key strength of the Pymetrics approach is how the tool is customised to each employer and job role, to avoid a one-size-fits-all definition of what employability looks like. Pymetrics works together with their clients to identify "top performers" among their staff in the role they are recruiting for. Ideally, such performance is based on objective criteria such as sales and revenues, rather than subjective manager and peer ratings. Top performers are then invited to take the Pymetrics test and their results are used to identify the soft skills most predictive of good performance in that role in the company. Not all jobs will require the same soft skills. For example, while it might be good for sales staff to be impulsive in their decision-making, analyst roles might be better off with more deliberative decision makers. Pymetrics can also help companies hire for new roles, in which case they use data from other companies that have hired for similar roles in the past.

Another crucial aspect of the Pymetrics approach is its custom validity testing. No model is put into use without extensive pre-deployment tests, which Pymetrics carries out in close collaboration with its clients. For example, the model's predictions are tested out-of-sample and the model is adjusted accordingly (through variable reweighting) by a multidisciplinary team (made up of data scientists and Industrial Organisation psychologists) using a supervised learning approach. Further testing can be done with data from workers outside the company and who have previously taken Pymetrics tests and agreed for their data to be used. This dataset contains several million observations on candidates from different industries, employers and geographies, and can provide a useful additional check on the model.

Based on some of these tests, Pymetrics claims that its tools can increase tenure (+198%) and sales (+28%), as well as decrease time to hire (-59%).

Source: OECD interview with Pymetrics.

There is some evidence that AI can help improve the quality of public and private employment services. For example, Cockx, Lechner and Bollens (2020[40]) have estimated that the use of a machine learning algorithm by the Flemish PES (VDAB) to reassign jobseekers to training programmes could result in jobseekers spending 20% less time in unemployment over a period of 30 months (as compared to the current assignment system), which means jobseekers are getting a better service. Similarly, Pôle Emploi uses AI to identify at an early stage which jobseekers will have the most difficulty getting back to work (Box 9). In the PrES sphere, an AI developer and business strategist interviewed as part of this project discussed an AI-based search tool that draws on historical data of successful matches to recommend candidate profiles to recruiting firms. The tool's performance was assessed according to its "precision", a metric reflecting the share of recommended candidates that a recruiting firm might be interested in. It was argued that a precision of 50% was fairly good (i.e., the recruiting firm would be interested in one out of two results returned). The person interviewed argued that AI had a higher level of precision (60%) than the average human being (30%)—even though a good and experienced search expert might still do better than the AI.

Box 9. Predicting length of unemployment spells using AI at Pôle Emploi (France)

A key task of caseworkers at Pôle Emploi is to identify, as soon as possible, the jobseekers who face most difficulty in returning to work and to establish with them a Personal Job Plan (*projet personnalisé d'accès à l'emploi* – PPAE) which specifies the actions the case worker and jobseeker will undertake in order to find work. At Pôle Emploi, all jobseekers have an obligation to draw up and follow a PPAE, however the content of these plans varies from one person to another, depending on the difficulty he/she faces in returning to work. Caseworkers therefore need to classify jobseekers in order prioritise their interventions. At Pôle Emploi, caseworkers are assisted in this task by an Artificial Intelligence tool which provides an estimated unemployment spell for each individual. Compared to previous statistical models, AI allows more refined predictions based on more information as well as a dynamic approach. AI does not only allow for a regular updating of the information used in these predictions, but it also captures data about the current and future state of the labour market ("weak signals"). For example, if a large firm will soon open a new branch in the region, such information can be taken into account as soon as it appears in the urban development plan.

Improving the candidate experience

AI could also be used to improve the candidate experience and, in turn, increase organisational attractiveness. To date, organisations adopting AI have often focused on achieving efficiencies, which has sometimes come at the detriment of candidate experience. This, in turn, can put off candidates and lower the applicant pool or even damage the reputation of the organisation (see Section 4). However, AI can be used to improve the candidate experience, notably by improving communication and greater personalisation of the hiring experience. An HR Innovation Strategist at a large consulting company interviewed as part of this project said: "There are two things candidates don't like: not getting any feedback on their application, and not knowing why a decision was made. Now, with AI, there is more emphasis on the candidate experience. Communication with candidates can be improved." AI chat bots can provide personalised feedback, and they are accessible outside working hours and can provide quick answers to candidate questions.

Addressing human bias

There is plenty of evidence that human decision-making can be biased (Krieger, 1995[41]) (Kang and Lane, 2010[42]) (Jost et al., 2009[43]) (Raymond, 2013[44]). For example, research has shown that White interviewers sit farther away from Black applicants than from White ones, smile less genuinely, and end interviews 25% sooner (Frith, 2015[45]). There is also evidence that people typically associate negative characteristics more strongly with disfavoured groups and that these negative associations can result in adverse decisions for the members of those groups, even when people believe they are acting fairly (Kim, 2017[46]). In economics, there is a large literature based on so-called "correspondence experiments" to measure hiring discrimination. In these studies, fictitious job applications, differing only in a randomly assigned characteristic, are sent in response to real job openings. By monitoring the subsequent call-back from employers, unequal treatment based on this characteristic can be identified. This literature has revealed discrimination against ethnic minorities, pregnant women and mothers, transgender people, Muslims, as well as people with disabilities (Baert, 2017[47]).

AI and hiring algorithms could help address human bias (Kim, 2017[46]) (Bornstein, 2018[48]). AI could introduce an element of objectivity and neutrality in the matching process. It can be more consistent than traditional assessment techniques, like in-person interviews, because it offers all candidates a standardised experience based on concrete data (van den Broek, Sergeeva and Husyman, 2019[49]) (Chamorro-Premuzic, 2019[50]) (Fisher and Howardson, 2022[51]). AI can also help address bias in other ways. Yam and Skorburg (2021[52]), for example, argue that algorithms can widen the pool of diverse job candidates through targeted advertising (see Section 2 for some examples). In 2018, LinkedIn introduced an AI feature that ensured that top search results seen by recruiters had a gender breakdown more representative of the potential applicant pool (Bogen and Rieke, 2018[20]). Joaquin Quinonero Candela from LinkedIn, interviewed as part of this project, said, "We aim to provide equal opportunities to equally qualified candidates, and we make sure our AI supports this goal." AI could also help to detect human bias and this can then be used to address it.

AI matching tools increasingly have built-in features to address bias. Garr and Jackson (2019[53]) show that "diversity and inclusion technology" is a booming business. They looked at 105 vendors in 2019 who, together, had a market size of approximately USD 100 million, with 40% of them experiencing more than 100% year-on-year revenue growth. Many vendors design their algorithms to remove bias (see Box 10 for the example of Pymetrics, and also Sánchez-Monedero, Dencik and Edwards (2020[23])). Several companies remove variables that are correlated with protected attributes, or give them less weight, when adverse impact is detected, with little effect on the predictive accuracy of their products (Raghavan et al., 2019[54]). In the United States, many of these tools are deliberately designed to shield companies from legal liability by complying with the 4/5 rule which requires that the likelihood of one group being selected should not be less than 80% that of any other group (WEF, 2020[39]).[8] Regarding the practice of reducing bias within PrES firms, an AI developer and business strategist interviewed as part of this project mentioned two approaches for reducing bias in their recommender tool. The first approach, known as "post-processing", involves setting quotas for the share of candidates recommended from each population sub-group. The second involves feeding the algorithm artificial profiles to train it and boost the chances of under-represented groups being selected. Where possible, the interviewees said that post-processing is the most effective method. However, in many use cases, the variable of interest may be unknown (e.g., gender). In this case, the AI must rely on proxies for the variable of interest, which are not always readily available. In the absence of proxy variables, the second approach of treating the data itself is an alternative.

Box 10. How Pymetrics addresses bias

Pymetrics helps companies improve diversity and avoid bias (disparate impact). There are various elements to this approach, starting with the building of the model, which avoids the use of soft skills which have been shown by the literature to be correlated with demographic characteristics. When training the model, Pymetrics asks its clients to provide as diverse a training sample as possible and, in the testing phase, measures that are shown to be biased are de-weighted until the model avoids disparate impact. In addition, Pymetrics helps its US-based clients comply with Title VII of the Civil Rights Act by providing evidence of the job-relatedness of the model, as well as documentation of the search for less biased alternatives.

Finally, there are some candidates who might have accessibility issues when carrying out a Pymetrics test, such as dyslexic candidates, those with attention deficit disorders, or those with colour blindness. For those candidates, Pymetrics makes accommodations to ensure that they are not disadvantaged by the test. Moreover, although candidates may request such an accommodation, the recruiting company will not know that the candidate declared having a disability during the recruitment process.

In one of its assessments, Pymetrics found that its tools increased in female representation (+62%).

[8] In reality, the 4/5 rule is only a guideline. The Uniform Guidelines on Employment Selection Procedures state that smaller differences can constitute adverse impact and greater differences may not, depending on circumstances (Barocas and Selbst, 2016[140]). The guidelines were issued by the five Federal agencies having primary responsibility for the enforcement of Federal equal employment opportunity laws, to establish a uniform Federal government position. The guidelines are designed to aid in the achievement of the goal of equal employment opportunity without discrimination on the grounds of race, colour, sex, religion or national origin. The guidelines apply to private and public employers, labour organisations, employment agencies, apprenticeship committees, licensing and certification boards, and contractors or subcontractors.

There is evidence that AI tools can help improve diversity. A managing director for workforce transformation at a large consulting company interviewed as part of this project said:

> *"AI can increase the diversity of applicant pools. There is evidence that women tend to self-select themselves out more than men do. A man might apply for a job if he meets 50% or less of the criteria, whereas women will only apply if they match a much higher percentage of the skills articulated in a role. With AI, because it automatically identifies opportunities that would be a good match, the percentage of women who complete applications has gone up and, in some cases, quite dramatically. Some organisations have seen a 40%, 50% or higher increase in the number of female applicants in their pool."*

Beyond anecdotal evidence, a study by Cowgill (2020[55]) showed that AI algorithms can increase the hiring of under-represented candidates, such as women, racial minorities, candidates without a job referral, graduates from non-elite colleges, and candidates with no prior work experience. Sühr, Hilgard and Lakkaraju (2020[56]) analyse gender bias in an online hiring platform. Simulating hiring scenarios with data from TaskRabbit, an online freelancing website, they show that fair ranking algorithms can improve the selection rates of women. However, the authors also find that the effectiveness of these algorithms is dampened in job contexts where employers have a persistent gender preference (e.g., moving assistance jobs), while their effectiveness improves the more the profiles of underrepresented candidates resemble those of the overrepresented group. Li, Raymond and Bergman (2021[57]) start from the premise that, in order to find the best workers over time, firms must balance "exploitation" (selecting from groups with proven track records) with "exploration" (selecting from under-represented groups to learn about quality). They argue that modern hiring algorithms based on "supervised learning" approaches are designed solely for exploitation. The authors build a CV screening algorithm that values exploration by evaluating candidates according to their statistical upside potential[9] and, using data from a professional services recruiting within a Fortune 500 firm, show that this approach improves the quality (as measured by eventual hiring rates) of candidates selected for an interview, while also increasing demographic diversity. Finally, Allred (2019[58]) looks at a pre-employment assessment of general cognitive ability (GCA) which tends to be a good predictor of job performance, but also suffers from very large racial-ethnic group differences and can therefore result in different selection rates for majority and minority group members. Allred (2019[58]) designs an algorithm which allows the use of GCA while minimising racial-ethnic group differences (however, the author finds that this comes at a cost of lower validity). In more qualitative research, HR professionals agree that AI can give a chance to atypical candidates who previously might not have made it through the screening process (Li et al., 2021[59]) – although that result is likely to depend on the kind of AI that is being used and how it is coded.

A human-machine collaboration might be the most effective way of addressing bias. None of the discussion above invalidates concerns that AI itself may perpetuate, exacerbate or even introduce new bias into the matching process (see Section 4). However, humans are also biased and there is evidence that people may hold AI to a higher standard than humans (Fisher and Howardson, 2022[51]) and that "the errors that we tolerate in humans become less tolerable when machines make them" (Dietvorst, Simmons and Massey, 2015[60]). People lose confidence more quickly in algorithmic than human forecasters after seeing them make the same mistake (Dietvorst, Simmons and Massey, 2015[60]). Such bias is an important barrier to adoption and could be very costly if the AI is actually better than humans at predicting. Fisher and Howardson (2022[51]) argue that people may be more willing to accept inconsistency in human decision

[9] In practice, Li, Raymond and Bergman (2021[57]) use an Upper Confidence Bound (UCB) contextual bandit algorithm: in contrast to supervised learning algorithms, which evaluate candidates based on their point estimates of hiring potential, a UCB contextual bandit selects applicants based on the upper bound of the confidence interval associated with those point estimates. That is, there is implicitly an "exploration bonus" that is increasing in the algorithm's degree of uncertainty about quality. Exploration bonuses will tend to be higher for groups of candidates who are underrepresented in the algorithm's training data because the model will have less precise estimates for these groups.

making because it can be linked to an actor's intentions, whereas such intentions are absent in the case of machines and algorithms. For instance, many papers and articles on this topic mention the example of Amazon's hiring algorithm which was found to be biased against women, whereas similar discrimination is a daily occurrence in human hiring decisions and does not receive the same media coverage. Moreover, a privacy director of a PrES interviewed as part of this project argued that the Amazon case is actually a good example of how to use AI in recruitment: Amazon developed AI based on their historical data, then they tested the tool, and they figured out there was a bias they had not previously been aware of. Subsequently, Amazon published these findings and dropped the tool, while going back to think about how to address the uncovered bias. The take-away from all this is that neither humans nor AI are perfect, but that a combination of their relative strengths might be the most effective way to reduce bias in matching.

4 Adoption Rates and Barriers to Further Adoption

The current use of AI tools in matching still appears limited and there a number of barriers to the greater use of these tools. On the one hand, there are barriers related to the readiness of organisations and people to use AI, including: a lack of skills, management culture, poorly prepared data and information systems, as well as resistance from staff who have to work with the tools. On the other hand, there are concerns about the technology itself and its limitations when it comes to matching which, following the OECD AI principles, can be subdivided into three broad categories: the robustness of the tools; infringements of human rights (including dehumanisation, privacy, and fairness); and challenges around transparency and explainability.

The adoption of AI for matching

The limited evidence on the adoption of AI tools suggests that current use in recruitment remains limited. Robust evidence on the rate of adoption of AI tools is lacking, and the existing estimates vary widely. It is also difficult to interpret many of these figures, because often little is known about the methodology and representativeness of these studies. In addition, studies are frequently not limited to AI, or take a broader look at AI in HR (as opposed to its use in recruitment only). Some estimates suggest that adoption is very high, at least among some types of firms. For example, 40% of HR functions of international companies from around the world say they are currently using AI applications (PWC, 2017[61]) and 88% of talent acquisition professionals say they use AI/big data in recruitment (Korn Ferry, 2018[62]). Another study found that 98% of Fortune 500 companies used Applicant Tracking Systems of some kind in their hiring process (Sánchez-Monedero, Dencik and Edwards, 2020[23]). Other estimates however, are considerably lower. A report from 2019 by the HR Research Institute estimated that only 10% of HR professionals in the United States made high/very high use of AI for talent acquisition, while 36% made low/moderate use of it. 38% said they did not use it at all (HR Research Institute, 2019[63]). A McKinsey (2020[64]) global survey found that fewer than 10% of companies used AI in human resources. Similarly, a survey of US HR professionals by CareerBuilder (2017[65]) suggested only 13% already saw AI as a regular part of HR. In a global survey, Bader (2019[66]) found that the use of AI in HR (15%) was less common than in other parts of the firm (22%). Finally, while not strictly about AI, Tambe, Cappelli and Yakubovich (2019[67]) found that only one in five firms had adopted analytics in human resources. Overall, this evidence seems to indicate that, apart from in large firms, the use of AI in recruitment is likely to be limited still.

Adoption among Public and Private Employment Services also appears low. In a survey of members of the European Public Employment Services (PES) Network, 12% of PES said they were using AI-based matching (European Commission and Pieterson, 2020[68])—although 76% said they were planning to use AI for matching, highlighting the growth prospects. Another survey of this network comes up with higher estimates, with 11 out of 19 PES saying they were using AI (in matching, as well as for other purposes) (Awol Group, 2019[69]) – although the authors warn that the difference between AI and other technologies was not always clear in this study. Interviews with large, international players in the PrES industry suggest that, although most of them appear to be using some sort of AI, adoption is still limited to cases where the technology is advanced and the risks are low (e.g., interview scheduling, communication with candidates/firms), and less so where risks are high and there are still challenges with the technology itself (e.g., matching).

Jobs websites like Indeed, LinkedIn, ZipRecruiter, Glassdoor, etc. all use AI to some degree. For example, LinkedIn uses AI and the wealth of user data that it has, to optimise their members' experience, including: recommending candidates to recruiters, suggesting jobs to individuals, and advising on connections to make.

Despite being in its infancy, the market for AI recruitment tools is expected to grow over the next few years. In 2019, the AI in recruitment market size was valued at $580 million worldwide, and was expected to have a compound annual growth rate of 6.8% over the period 2020-2025 (IndustryARC, n.d.[70]). Whether this growth materialises will depend on whether some of the barriers to further adoption can be overcome. These barriers are discussed in the next few sections of the paper.

Barriers related to organisation and people readiness

There are many non-technological barriers to the adoption of AI for labour market matching. A study undertaken by Deloitte found that, although 75% of companies believed that using human capital analytics was important for business performance, only 8% viewed their organisational capabilities in this area as "strong" (Minbaeva, 2018[71]). Similarly, among PES, the majority (80%) say the use of analytics would help improve the outcomes of labour market interventions and programmes, however they also claim not to have the right resources to use analytics (Accenture, 2015[72]). The following sub-sections explore organisations' and people's readiness for the use of AI tools in matching, including: the skills of the staff interacting with these tools; the lack of preparedness of data and information systems; and the attitudes of both management and staff.

Lack of skills

The successful adoption of AI requires staff with the right skills to work with, or alongside, AI. While workers do not necessarily require technical AI skills, there does seem to be a need for analytical capabilities to work with AI and big data, even if it is just understanding the principles of data-based analysis, including its power and limitations (Levenson, 2005[73]). There is evidence that HR professionals who have stronger analytical skills (and are therefore in a better position to analyse and interpret data) perform better at their job (Kryscynski et al., 2018[74]), and that there are four reasons for this: (i) analytical skills allow them to make better decisions; (ii) insights from data can be leveraged to effectuate change; (iii) the ability to use and interpret data and information allows HR professionals to discover new insights; and (iv) analytical skills make HR professionals better at communicating and coordinating with other numbers-driven functions, such as R&D, sales, finance, etc.

However, analytical skills appear to be lacking among many recruitment professionals. This is true for HR professionals (Giermindl et al., 2021[75]) (Dahlbom et al., 2020[76]), as well as staff in PES. For example, even though 80% of PES covered in a study by Accenture (2015[72]) said that the use of "analytics" would help them improve services, they also mentioned that they did not have the right resources to do so. In particular, 78% said they needed more employees skilled in analytics. This latter point is echoed by Angrave et al. (2016[77]) who, from their discussions with HR professionals, concluded that even organisations that have invested in such technologies still lack many analytical skills. This contrasts with other areas in the organisation that depend on data analysis (such as R&D and finance), where there are often skilled analysts (Lawler, Levenson and Boudreau, 2004[78]). This may partly explain why some HR staff are concerned about AI. In a study by Bartsch (2020[79]), 71% of HR representatives interviewed said that they were personally not very well prepared for the use and deployment of AI. This is one of the reasons the Flemish PES, VDAB, invests heavily in training individuals to interact with its AI systems (Box 11).

Box 11. How VDAB, the Flemish Public Employment Service, trains caseworkers and jobseekers in the use of AI

The Flemish PES, VDAB, is one of the front-runners when it comes to the use of AI in delivering PES services. Many of its employees and customers have been using AI applications for years and have gained some understanding of the benefits, risks, and limitations associated with AI. This notwithstanding, VDAB continues to provide ongoing support and training to help individuals make effective use of AI. VDAB also takes steps to assess and develop the core skills needed to make good use of data-driven tools, including AI applications. For example, a study was launched in 2021 to gauge the level of data literacy of VDAB employees, with the aim of providing recommendations to the management team.

VDAB does not provide tool-specific training. All training for AI applications is integrated into a broader curriculum that focuses on the various missions entrusted to caseworkers. Caseworkers learn how they should use a given AI application as part of the bundle of skills they require to deliver the service(s) to which the application is attached.

Caseworkers also have access to a variety of support materials and events that enable them to develop their general knowledge about AI on an ongoing basis, as well as to find answers to the questions they may have on any AI application in use within the PES. To this end, VDAB relies on various pedagogical approaches, such as:

- Information sessions and events (e.g., “Coffee with the Future”, “Lunch & Learn sessions”, “Digiwijs”) that allow for open communication between caseworkers and IT people on a variety of topics, ranging from digital literacy to general features of AI, as well as more specific questions concerning the AI applications caseworkers use in their daily work. In addition, caseworkers can gain a deeper insight into AI solutions when they are involved in the development process to share their field experience and expertise with developers, thereby participating in various innovation labs, working groups and testing procedures.
- Online courses and webinars that, in some cases, are accessible to caseworkers and customers alike. VDAB also organises “Inspiration” sessions which look at AI applications from different perspectives. They typically bring together caseworkers, managers and IT people from the PES, as well as VDAB customers who may use AI applications in self-service mode or indirectly through their PES counsellor. All tips, examples and information gathered during these inspiration sessions can in turn serve as a basis to develop practical guidance and support materials available to everyone.

Data and information systems

Besides skills, the other ingredient often lacking to successfully implement AI is a good data and information system. In their survey of firms, Lismont et al. (2017[80]) found that the most common challenge for analytics boiled down to data management issues, and Lawler, Levenson and Boudreau (2004[78]) found that organisations were more optimistic about having the skills than about having the data required for analytics.

Data held by organisations is frequently of poor quality and unorganised. Dahlbom et al. (2020[76]) argue that the use of big data to enrich analytical insight is still a “futuristic vision” because most organisations have a lot of work to do to systematise and automate even the most basic reporting. Data is often spread across many different HR systems and few firms have a unified, integrated and cleansed database (Giermindl et al., 2021[75]). The information that could be useful for AI often has to be extracted from multiple databases, converted to a common format, and joined together before it can be used (Tambe, Cappelli and Yakubovich, 2019[67]). The same problem is identified by PES, 68% of whom say that they need better data to fully benefit from analytics (Accenture, 2015[72]).

The data available to recruiters is often too focused on the HR function and on measures of efficiency (e.g. cost of recruitment, time to hire, etc.), rather than on outcomes for the firm/organisation as a whole. The real value of AI in matching depends on the ability to link HR data to performance data, so that characteristics of candidates can be linked to future outcomes for the firm (Lawler, Levenson and Boudreau, 2004[78]). The challenge, therefore, is to link data on recruitment to data from finance, operations, and other parts of the organisation. As argued by Cappelli (2017[81]), "Unless we can get the data in those two databases to be compatible, there is no way to ask even the most basic questions, such as which applicant attributes predict who will be a good performer." In pratice, most firms do not even know what types of data are available to them or in what form (Minbaeva, 2018[71]). Moreover, data is often collected over different time periods and organisational levels (e.g., individuals, teams, departments, business units), and there tend to be breaks in time-series as well as inconsistencies (Minbaeva, 2018[71]). In some cases, there might be a need to change data collection systems, and collect new types of data.

Small firms are unlikely to have the data required for AI (Nocker and Sena, 2019[82]). Recruitments are relatively rare occurrences, particularly in small firms. Yet the whole point of machine learning and other data science techniques is that they require large numbers of observations, which raises the question of whether many of these tools can only be used by large organisations. Instead of using their own data, small firms would have to settle for solutions that are based on other companies' data and may, therefore, not necessarily be a good fit. Cappelli (2017[81]) argues that, in most companies, "HR does not actually have big data", which makes him question the usefulness of AI for these organisations: "There is almost no reason for HR to use the special software and tools associated with big data. For most companies, the challenge in HR is simply to use data at all […] they need simple software—sometimes even Excel spreadsheets can do the analyses that most HR departments need."[10]

Resistance from management and/or staff

Challenges to adoption of AI can come from the very top of the organisation as managers either lack an understanding of these technologies, or fail to see the benefits of investing in them and perceive a weak business case for change (Minbaeva, 2018[71]). Sometimes, such investments are seen too much from the HR function's internal perspective alone, thus missing the wider business case (Dahlbom et al., 2020[76]). More generally, a lack of clear strategy has been identified as the top challenge to the adoption of AI in organisations (McKinsey, 2018[83]), while fear of change may also be partly to blame (Fraij and Lászlo, 2021[84]). A managing director for workforce transformation at a large consulting company interviewed as part of this project said that "technology fatigue" may be another factor:

> *"There has been a tremendous influx of investments and some organisations are feeling overwhelmed by the process of assessing which ones to work with, the risks and the benefits, and the whole change management journey. […] Policy should spend more time on educating people and companies about the positive sides of AI and what it can do for society. People need to know what AI is and how it works. We need stronger use cases to tell stories about where AI is being used, why organisations use it, the process that led to*

[10] An additional challenge is the need to have data protection officers / employees with such skills in order to process the data in compliance with existing regulations.

validation, so that organisations don't need to reinvent the wheel every time they are considering a new technology."

Some staff in human resources and/or employment services may oppose the use of AI for fear of losing their jobs. In recent years, AI has made significant progress in areas like information ordering, perceptual speed, speech recognition, written and oral comprehension (Georgieff and Hyee, 2021[85]). This means that many tasks involved in matching, such as sifting through CVs, checking answers to screening questions, responding to basic candidate queries, scheduling interviews, etc. could be automated by AI. Workers involved in matching may therefore worry about the threat that AI technologies pose to their jobs (Hekkala and Hekkala, 2021[34]). In one study, more than a quarter of HR professionals thought AI in the workplace would be disrupting and/or displace jobs in the next 10 years, while only 21% said they were excited about AI (Allegis Group, 2017[86]). Such concerns could lead to lower technology acceptance and trust, and decreased intention to use the technology (Ulfert, Antoni and Ellwart, 2022[87]). They could also hamper the adoption of new technologies if managers receive pushback from workers (Li et al., 2021[59]).

That being said, many staff see AI as complementing, rather than substituting, their role. In practice, automation is unlikely to lead to significant job loss among HR professionals, and PES and PrES caseworkers. The evidence for the labour market overall suggests that, over the last couple of decades, very few occupations have experienced negative employment growth, despite the adoption of automating technologies (Georgieff and Milanez, 2021[88]). This is partly because new technologies tend to make workers more productive which, under the right conditions, results in a greater need for workers (Lane and Saint-Martin, 2021[89]). Lisa and Talla Simo (2021[90]) find that recruiters do not necessarily see AI as a threat to human jobs, but rather as playing a complementary role. To some extent, what to automate and what not to automate will be a choice. Professionals in charge of matching prefer not to automate everything, but to maintain some human elements in the process (Mirowska and Mesnet, 2021[91]). Laurim et al. (2021[92]) show that, as long as final decision-making power is still with the human being, recruiters see no risks of AI replacing their jobs. Similarly, HR professionals interviewed by Vanmeirhaeghe (2021[93]) appear positive about a future with AI, as long as a human touch is kept and the human remains in control.

AI could have a positive impact on the job quality of workers involved in matching. AI offers opportunities for reducing the time that recruiters and PES and PrES staff spend on the more repetitive and routine aspects of their jobs. For example, in France, the PES Pôle Emploi, introduced an AI tool that partly automates the responses to e-mails received by caseworkers—which can be both numerous and repetitive. As a result, workers could spend more time on tasks where human creativity and emotion are more important (e.g. interviewing, rapport building and negotiating) (Johansson and Herranen, 2019[94]) (Upadhyay and Khandelwal, 2018[95]), and on tasks with higher value-added (e.g. recruiting for more specialised hires) (Li et al., 2021[59]). This, in turn, could imply an improvement in the job quality of these workers. As one Workforce Strategy & Transformation Leader at a large consulting company said: "While AI can automate some work, it will also create new roles for staff, like making sure the bots are trained and that the technology is monitored, that the results are tracked, and that there is no bias. This will require an additional layer of different skills. The recruiter's job will not disappear, it will change, it will be enhanced."

However, AI also presents some risks with regards to job quality. The shift to higher value-added tasks could require new skills and add to job strain. The interaction with automated decision-making tools might also result in a loss of autonomy for HR professionals and caseworkers.[11] As a system becomes more autonomous, and its reliability and robustness increase, workers tend to lose situational awareness and become less likely to be able to take over manual control when needed (Endsley, 2017[96]). While workplace monitoring is not new, AI greatly expands the scope and possibilities for tracking worker performance, and there is evidence that there are psychological risks attached to such monitoring,

[11] A loss of autonomy need not necessarily be a bad thing. Indeed, there is some evidence that lower human autonomy could lead to lower "technostress" and decreases information load (Ulfert, Antoni and Ellwart, 2022[87]).

including stress and anxiety, which can result in a reduction in worker motivation and undermine trust between employers and workers (Moore, 2020[97]). Although the evidence is not clear-cut, there are indications that, under certain circumstances, increased surveillance could be counterproductive even for productivity and result in increased absences from work and an atmosphere of hostility in the workplace (Sarpong and Rees, 2014[98]).

Apart from the impact on jobs, workers interacting with AI may have concerns about the technology itself. These technology-related concerns will be discussed in further detail in the next sub-section of the report. However, as an HR Adviser in a large manufacturing firm interviewed as part of this project said: “There is some resistance to the adoption of these tools from more traditional HR practitioners.” Such resistance may not necessarily be a bad thing. Indeed, the same HR Adviser added, “Before you adopt AI, you need to understand the risks, as well as how it works. Many AI tools are black boxes.” To some extent, the training of staff could help with the acceptance of AI tools in the workplace, provided that these tools are trustworthy in the first place. As an HR Innovation Strategist at a large consulting company interviewed as part of this project said: “Helping people understand what AI is, what it can do, and how it works can help overcome barriers to adoption and improve trust.”

The success of AI in matching might also partly depend on whether workers are involved in the design, implementation and evaluation of AI tools. As Bodie et al. (2017[22]) point out, “it is critical for employees to understand the ongoing processes, rather than feeling like a set of test subjects. Employees need to have a voice when it comes to implementing these practices”. Similarly, Giermindl et al. (2021[75]) argue that workers should be given “the freedom to speak out if people analytics’ decision-making recommendations go against their own insight, thus promoting an open culture of exchange and error”. OECD work has shown that representative workers’ voice is associated with mitigated risks of AI on employment, working conditions and wages (Krämer and Cazes, 2022[2]) Worker voice can be heard through informal consultations in the workplace, but also through more formal mechanisms, such as collective bargaining. In France, all decisions with regards to the development and implementation of AI within the PES, Pôle Emploi, are taken within the context of the triannual strategic plans which are first discussed at a tripartite committee. For every AI project implemented, the workers who are likely to be affected are consulted on both the desirability and feasibility of the project.

Barriers related to concerns about the technology

There appears to be significant concern about the use of AI in recruitment. In a survey of the American public, just 3% said they were “very enthusiastic” about job candidates being vetted by computer programmes, and 28% were “not at all enthusiastic”. Meanwhile, a sizeable majority of Americans (76%) said they would not want to apply for jobs that used a computer programme to make hiring decisions. The vast majority thought that computer programmes would do worse than humans at various aspects of the hiring process, including: hiring candidates with diverse backgrounds, hiring qualified candidates, hiring candidates with non-traditional work experience, and hiring candidates that fit well with the company’s culture (Smith and Anderson, 2017[99]). In their review of the literature, Langer and Landers (2021[100]) found that automated hiring systems predominantly resulted in negative applicant reactions. In a study by the American Staffing Association (2016[101]), 77% of jobseekers said that they preferred human interaction throughout the course of the job hunt. Laurim et al (2021[92]) found that candidates also reacted negatively towards chat bots. Similarly, a privacy director of a PrES interviewed as part of this project said that jobseekers in some countries did not feel comfortable using a chat bot. Concerns are not limited to individuals—many companies too express doubts. An HR Adviser in a large manufacturing firm interviewed as part of this study said: “As a company, we have looked at various AI solutions but, so far, we have never really had a good feeling about them.”

The perceived risks of AI in matching depend on the stage of the matching process it is used in. As one HR professional put it, "AI software tools work the best when they are implemented in the beginning phase of the recruitment process, where there's more administrative tasks and not a need for humans" (Johansson and Herranen, 2019[94]). In private employment service firms, this involved tools to identify job vacancies, to optimise the language used in job descriptions, and to target job advertisements to specific groups. Similarly, a Workforce Strategy & Transformation Leader at a large consulting company interviewed as part of this project said: "We see most adoption at the screening and interview scheduling stages, because those involve really time-intensive and manual work." Hekkala and Hekkala (2021[34]) argue that AI cannot replace humans in recruitment phases like building relationships, interviewing and, more generally, those where emotional intelligence is required. Laurim et al (2021[92]) report that recruiters are positive about the use of AI in the creation of job advertisements, but less so in the evaluation and selection of candidates. Wehner, Köchling and Warkocz (2021[102]) find that candidates also experience AI more positively in the early stages of the selection process than in the later ones, and Smith and Anderson (2017[99]) find that the general public would feel better about hiring algorithms if they were used only for the initial vetting of candidates.

The recruitment phases where people feel more comfortable about the use of AI tend to be those where the stakes are lower. One of the reasons people worry about the use of AI in matching is that the stakes can be high. AI is not being used simply to make a movie recommendation or target an online advertisement for shoes. It is used to make decisions about who sees a job advertisement, who passes the screening process and gets an interview, and even who passes that interview and the terms and conditions of the job offer made in the end. These are decisions that have an important impact on people's careers. Langer, König and Papathanasiou (2019[103]) show that individuals react differently to the use of AI depending upon whether the context is high- or low-stake. The limited data on adoption of AI in HR confirms that companies are also more reticent to use automation in later stages of recruitment. A HR Adviser in a large manufacturing firm interviewed as part of this project said: "Using AI for core recruitment tasks seems unlikely at this stage, especially automatic decision-making. However, the use of AI for more peripheral tasks could be helpful, for example chat bots to respond to basic candidate queries. Or for sourcing, because you are very far from the final decision process."

Negative attitudes towards automated decision-making affect candidates' perception of the recruiter and could result in smaller applicant pools. There is evidence that the characteristics of the recruitment process affect applicant attraction to a particular job (Uggerslev, Fassina and Kraichy, 2012[104]). Job applicants tend to trust organisations less if these organisations use AI/ML to make decisions (Gonzalez et al., 2019[105]). Langer et al. (2020[106]) found that automated interviews had a negative impact on organisational attractiveness, Will, Krpan and Lordan (2022[107]) conclude that applicants view organisations deploying AI in hiring as less attractive than those hiring through humans, and Acikgoz et al. (2020[108]) not only found lower organisational attractiveness, but also decreased job pursuit and stronger litigation intentions when automation decisions were used in interviews. Laurim et al. (2021[92]) found that recruiters were hesitant to use video applications (i.e. applications that analyse videos of candidates and create personality profiles by registering choice of words, articulation, or facial expressions and gestures) because they did not want to discourage applicants from applying. Similarly, Vanmeirhaeghe (2021[93]) argued that candidates did not like chat bots, and that they could harm their view of the firm.

Following the OECD AI principles, concerns around AI technologies for matching can be subdivided into three categories: those related to the robustness of the technology; concerns around human rights; and questions around transparency and explainability. The next few sub-sections will discuss each of these in turn as they apply to the domain of matching. For a broader discussion of the ethical issues surrounding the use of AI in the workplace, the reader is referred to (Salvi del Pero, Wyckoffi and Vourc'h (2022[1]).

Robustness

The robustness of AI is a critical factor in building trust in such tools. At the moment, a perceived lack of reliability of AI tools is still one of the reasons people tend to prefer humans to remain in control of the recruitment process, however there is evidence that humans start to prefer AI control when its accuracy gets better (Langer and Landers, 2021[100]).

In some tasks, AI may be as good as or better than humans. In an experiment conducted in an online labour market (Upwork), Horton (Horton, 2017[109]) found that algorithmically recommending workers to employers substantially increased hiring, because it was cheap for employers to act upon these recommendations and the recommended candidates were similar to the kind of workers the employers would have recruited themselves. In another study, Cowgill (2020[55]) used data from a large company that trained an algorithm to predict which candidates would pass its interviews. An algorithm was then allowed to randomly override the choices made by human screeners. The results showed that the candidates picked by the machine were 14% more likely to pass interviews and receive a job offer; 18% more likely to accept job offers when extended; as well as more productive once hired as employees. However, robust studies like these remain rare and the tasks they focus on are limited.

Concerns around the reliability of many tools remain, especially facial and voice recognition tools, many of which are used in automated video interviews for example. Some AI applications are being sold on the grounds that voice, facial expressions and body movement are predictors of personality and/or emotion, and that AI can help interpret these in a more objective way (Biel, Teijeiro-Mosquera and Gatica-Perez, 2012[110]; Nguyen and Gatica-Perez, 2016[111]). However, several studies have shown that commercially available recognition tools may be unreliable. A study by researchers at UCL concluded that, when it comes to reading emotions on people's faces, AI still lags behind human observers, in particular when facial expressions are spontaneous rather than posed (Dupré et al., 2020[112]). These tools are also less reliable for some population sub-groups than for others. Rhue (2018[113]) finds that facial recognition tools consistently interpret Black faces as angrier than White ones, and Raji and Buolamwini (2019[114]) found larger error rates for darker-skinned females in commercial facial analysis models. Accuracy might also vary by age: facial recognition tools are often trained on mature faces, so they may struggle with younger ones (Howard, Zhang and Horvitz, 2017[115]). Further concerns have been raised about facial and voice recognition software discriminating against people who have disabilities such as deafness, blindness, speech disorders, and people who have survived a stroke (Fruchterman and Mellea, 2018[116]; Guo et al., 2019[117]). Speech recognition systems have also been shown to struggle with dialects (Tatman, 2017[118]) and non-native speakers, and tend to perform worse for women (Tatman, 2017[118]). In turn, concerns around reliability translate into lower adoption of such tools. A Workforce Strategy & Transformation Leader at a large consulting company interviewed as part of this project said: "The lowest adoption is in facial recognition. Those tools are available, but there are lots of questions around bias as well as legal risk."

The theoretical underpinnings of some AI tools are shaky. Returning to some of the examples mentioned above, there are questions about whether inferring emotions from people's facial expressions and voice may even be possible. NYU's AI Now Institute has called HireVue's facial and voice analysis tools "pseudoscience" (Harwell, 2019[119]) and Barrett et al. (2019[120]) argue that how people communicate anger, disgust, fear, happiness, sadness, and surprise, varies substantially across cultures and situations. In response to these concerns, Microsoft has recently stopped offering emotion recognition as an option in its products (DeGeurin, 2022[121]). One interviewee for this project said, "Why would someone's hand gestures in a video indicate whether they are a good project manager or not?" The theoretical underpinning some other types of AI tools have also been questioned. Some tools track digital footprints (e.g., "likes" on Facebook and/or language used in e-mails and social media posts) to try and predict personality traits and values. It is easy to see the attraction of these tools, since they offer a faster and cheaper alternative to traditional selection tests and they do not require jobseekers to complete any assessments. However, while some researchers have claimed that such tools can be accurate (Kosinski, Stillwell and Graepel, 2013[122]; Schwartz et al., 2013[123]; Garcia and Sikström, 2014[124]; Boyd et al., 2015[125]; Azucar, Marengo and Settanni, 2018[126]; Kern et al., 2019[127]; Mori and Haruno, 2021[128]), others have questioned the relevance of their predictions for job performance and/or whether they add anything over more traditional predictors. For example, some AI tools carry out social media background checks by looking for toxic online behaviour, such as the use of foul language. However, it is not clear that online behaviour is relevant to a person's professional activity (Bogen and Rieke, 2018[20]). Moreover, definitions of what constitutes "toxic" or "concerning" content are often vague and highly subjective (Duarte, Llanso and Loup, 2017[129]). Van Iddekinge et al. (2013[130]) found no correlation between job performance and turnover, on the one hand, and social media profiles, on the other. They concluded that social media did not contribute to the prediction of these outcomes beyond more traditional predictors, such as cognitive ability, self-efficacy, and personality. This may be partly because people portray themselves positively in social media posts and their online presence may therefore be no more genuine than their CV (Chamorro-Premuzic et al., 2016[21]). Another application of AI in recruitment is the gamification of assessment tools, which is based on the assumption that people's behaviour in games can reveal aspects of their personality (Ewell et al., 2016[131]), and is a test of intelligence (Foroughi et al., 2016[132]; Ángeles Quiroga et al., 2015[133]; Unsworth et al., 2015[134]). In practice, however, there is still little evidence to back these claims up (Chamorro-Premuzic et al., 2017[135]). At the same time, these concerns are not new and have also been raised in the past with non-AI recruitment tools. As one interviewee for this project mentioned, "In the 1970s and 1980s we had people analysing candidates' handwriting for hints about their personality, yet there was little scientific evidence to show that such approaches were valid."

The quality of AI tools depends to a large extent on the quality of data used to train them—an issue which, in machine learning circles, is sometimes referred to as "garbage in, garbage out". In the recruitment context, a key objective of AI tools is to predict future worker performance. However, it is difficult to see how AI could be a reliable predictor of future performance when existing measures of worker performance are so poor. 86% of Fortune 500 executives admit that their companies do not know who are high and low performers (Keller, 2017[136]) and evaluations of worker performance tend to be very subjective and suffer from various sources of bias (Cecchi-Dimeglio, 2017[137]; Stauffer and Buckley, 2005[138]; Kraiger and Ford, 1985[139]).[12] Worse: instead of performance data, many AI tools used in the matching sphere simply use data like who has been successful at obtaining a job in the past or, at best, the tenure of such workers—all of which are very poor indicators of job performance. As an HR Innovation Strategist at a large consulting company said, even the most common "sources of data used, like job descriptions and resumes, are not actually of great quality and only poorly capture what people can do."

Algorithms tend to oversimplify reality and, as a result, risk misrepresenting it. As Tambe, Cappelli and Yakubovich (2019[67]) point out, "it is not easy to measure what constitutes a good employee" and "not all details of operations leave digital traces that can be captured, nor can all traces left be extracted and converted to a usable format at a reasonable cost". As a result, AI tools tend to define job performance in terms of a few measurable characteristics only (e.g. sales, absences from work, tenure) and use a limited set of measurable dimensions to predict them (Barocas and Selbst, 2016[140]).[13] This is only a very partial reflection of what success on the job looks like and helps explain concerns among the general public that AI cannot capture the true nature of an applicant and that it would overlook many important intangible qualities (Smith and Anderson, 2017[99]). Similarly, many HR professionals are sceptical of AI because they

[12] In addition, performance may not be the only criteria that organisations are looking for in a new recruit. Organisations may be willing to trade some performance for having someone who is pleasant to work with. Such characteristics may be even harder to measure.

[13] Some tools might only look at who has been hired in the past, however there is no guarantee that just because someone has been hired, that they will necessarily be good at the job. The real goal of hiring tools should be to predict performance on the job.

question whether people can be reduced to metrics (Angrave et al., 2016[77])—a practice sometimes referred to as "reductionism".

Reductionism could lead to reverse engineering which, in turn, could undermine the reliability of AI tools. Reverse engineering refers to the process through which one attempts to understand, by using deductive reasoning, how an AI tool reaches a certain decision/prediction. Knowledge gained this way can subsequently be used to try to mislead or play the AI. In the platform economy, where algorithmic management is common, examples abound of how workers have tried to make sense of the algorithm, shared their knowledge through online forums, and subsequently used the rules to their own advantage (Kinder, Jarrahi and Sutherland, 2019[141]; Lee et al., 2015[142]; Mohlmann and Zalmanson, 2017[143]). When dealing with AI in recruitment, candidates have been advised to copy texts from job advertisements and shrink the personal statement in order to trick the algorithm (Thomé, n.d.[144]). One interviewee as part of this project mentioned that some individuals have been found to add fictitious information (e.g., having a PhD from a reputable university) in white font to their CV, so that the AI would be able to detect the information, but not the human eye. Universities are also beginning to offer training to students to prepare for AI-based interviews (Burke, 2019[145]).

AI risks putting recruitment into a straitjacket. Algorithms put people into boxes or to label them. Candidates deplore the fact that, when writing their CVs, they are better off using generic terms that the AI will recognise, than expressing their own personality (Hekkala and Hekkala, 2021[34]). A related danger is that, over time, vacancies across companies will be increasingly written in the same way (Vanmeirhaeghe, 2021[93]). Just as candidate personality becomes less important, so does company culture/image. For example, because AI feeds on historical data on existing roles within companies, it might not be very helpful in cases where companies are interested in hiring an entirely new profile (Vanmeirhaeghe, 2021[93]). Moreover, as one HR Adviser in a large manufacturing firm said: "Often, you select candidates not on the basis of their past performance, but for their potential, their capacity to learn. This is not something that AI can help with." Concerns have also been voiced that AI may not be good at dealing with cultural differences in hiring (Li et al., 2021[59])—something which might contribute to bias and discrimination in hiring (see below). Some commentators fear that the use of (the same) algorithms could lead to "monoculture"—i.e. choices and preferences of employers becoming homogeneous in the face of algorithmic curation—which could not only be bad for individual job candidates who do not tick the right boxes, but could also, under certain circumstances, reduce the quality of matches overall (Kleinberg and Raghavan, 2021[146]). An HR Innovation Strategist at a large consulting company said:

> *"The recruiter really is the expert when it comes to hiring, and you want that expert oversight to validate AI recommendations. By design, AI solutions will fit candidates to the distribution – eliminating the tails on either end. My best employees were most often one-of-a-kind talents. If you rely solely on AI selection tools, you may miss out on the candidates that are so good they are unlike anyone else."*

Human-centred values and fairness: Dignity

Another concern with the use of AI in matching, is that it might dehumanise the matching process and put off potential candidates. AI is an automating technology and, as with many of these technologies, there are concerns that, by replacing humans, it will dehumanise recruitment, which has traditionally heavily relied on personal interaction (HR Research Institute, 2019[63]; Fritts and Cabrera, 2021[147]). This risk has arisen partly from the fact that organisations developing and adopting AI have been very focused on achieving efficiencies, and less so on improving candidate experience (Vanmeirhaeghe, 2021[93]). Dehumanisation would matter in particular if it impacted negatively on candidate experience and, as a result, on organisational attractiveness, since this would result in candidates withdrawing their applications and reducing the pool of potentially suitable applicants available to the hiring firm. This, in turn, could reduce the quality of matching.

Negative reactions to the use of AI in matching appear to be linked to the kind of skills that applicants believe are needed in the recruitment process, which are predominantly "human" skills. Lee (2018[148]) found that individuals were fine with algorithmic decisions for mechanical tasks (such as scheduling), but expressed negative emotions for tasks that they considered to require human skills (such as hiring). This negative reaction was attributed to the dehumanising experience of being evaluated by machines. Similarly, Gonzalez et al. (2019[105]) found that participants reacted less favourably to AI/ML decision makers than to human ones, citing concerns about the lack of dignified treatment and communication. Langer et al (2020[106]) confirm that applicants feel less social presence in highly automated interviews. Roheel Ahmad, managing partner at executive recruiter Forsyth Barnes believes that "The biggest factor going against greater use of AI is its lack of emotional intelligence. […] As advanced as AI can be, the human element allows subjectivity for individual cases. Career choices are very personal, and the opportunities we present to people are life changing. It takes a real understanding of another person and being flexible on an individual's circumstances to know what's best for them or where their skills are best utilised" (Morrison, 2022[149]).

Human rights: Fairness, bias and discrimination

In the context of employment, AI technologies have received a lot of negative publicity as a result of biased recommendations or decisions taken by some tools. One of the most famous examples is when Amazon's automated hiring tool excluded women because it had been trained on historical hiring data, which consisted primarily of male candidates (Giermindl et al., 2021[75]; Dastin, 2018[150]). Examples abound also of algorithms that exclude certain population subgroups from targeted job advertisements. Google, for example, was found to advertise highly-paid jobs less frequently to women (Datta, Tschantz and Datta, 2015[151]), and Lambrecht and Tucker (2019[152]) found that STEM jobs were less likely to be shown to women.

Bias is not only harmful to candidates, but also to employers. Employers face reputational and legal risks if the tools they use are found to be discriminatory. In addition, using biased tools means that they could miss out on scarce talent in the labour market. Fuller et al. (2021[153]) argue that automated recruitment systems, and the bias that is inherent in them, are a key reason why companies struggle to find talent despite an abundance of "hidden workers". Moreover, as Ajunwa (2016[154]) points out, even though human beings are also biased, the adverse impact of biased AI could be far greater by virtue of the volume and the velocity with which automated hiring tools make decisions, so any bias could be magnified and multiplied.

A first potential source of bias in AI tools can arise when the quality and the representativeness of data used to train AI vary in ways that correlate with certain group membership (Barocas and Selbst, 2016[140]). The quality of data for members of a certain sub-group may be poorer (e.g., less accurate, timely, or incomplete), or the data may be less representative of that particular sub-group in the general population. This may be because these groups have lower access to the internet, or have historically been less likely to apply for certain jobs. Lerman (2013[155]) writes that certain people "live on big data's margins" and that they are less "datafied" than the general population. Crawford (2013[156]) argues that "[b]ecause not all data is created or even collected equally, there are 'signal problems' in big-data sets—dark zones or shadows where some citizens and communities are overlooked or underrepresented." AI based on such data may result in skewed, or less accurate, conclusions for these groups. As an HR Innovation Strategist at a large consulting company said: "Because the training datasets that are used to train them don't have a diverse set of images to work from, facial recognition systems are unreliable and often discriminatory when applied across a broad population." Moreover, if AI presents accessibility barriers for certain individuals (e.g. those with limited internet access or poor digital skills, or individuals living with disabilities[14]), then this could reinforce issues around the representativeness of data and potential bias creeping into the system (Bogen and Rieke, 2018[20]).[15]

A second potential source of bias in AI algorithms can arise if the training data reflects past bias in human decision-making. As Raub (2018[157]) puts it, "Algorithms are in large part our opinions embedded in code". As mentioned in Section 3, human beings are often biased and, if AI is trained on historical data that reflects biased human decisions, then it will tend to replicate, and possibly exacerbate, that bias. Such bias can be deeply engrained in society and can affect, for example, the kinds of jobs that men and women apply to, or what gender employers prefer for certain jobs.

A third source of possible bias relates to the building of the models themselves—a bias which is sometimes referred to as "technical bias". It is a mistake to think that data mining models are "neutral" (Kim, 2017[46]). Developers need to make choices about the outcomes to be predicted, what data will be used to train the AI, how data will be labelled, what variables will be included, etc. These choices are not innocuous and could introduce bias into the algorithm because they reflect the world views of the developers and/or the socio-economic context in which they are made (van Es, Everts and Muis, 2021[158]). As Faraj, Pachidi and Sayegh (2018[159]) put it, "algorithms are imbued with the value choices of their designers, whether these have been made implicitly or explicitly". The lack of diversity among AI developers may exacerbate such bias. For example, Simonite (2018[160]) estimated that just 12% of machine learning researchers were women.

Simply removing "protected" characteristics (such as race and gender) from a model does not necessarily make it bias-free. In many countries, it is illegal to select on protected characteristics, like race, ethnicity or gender. Using such information as a basis for recruitment decisions would violate anti-discrimination law. However, even when the employer has no intention to discriminate, and all such characteristics have been removed from the data, bias and discrimination could still occur if the remaining characteristics are correlated with these sensitive characteristics as well as with the outcome of interest (Barocas and Selbst, 2016[140]). Where the use of such "neutral" but correlated variables is intentional, this is referred to as "masking" (Bodie et al., 2017[22]) or "data-laundering" (Ajunwa, 2016[154]) and some have argued that the advent of big data and AI makes these practices easier to apply, while also more difficult to detect (Barocas and Selbst, 2016[140]). Such biased decisions end up being "buried within a black box" (Sánchez-Monedero, Dencik and Edwards, 2020[23]). One of the most famous examples of this type of bias, is where an employer wanted to maximise worker tenure and found that commuting distance was the single most important variable that predicted how long a worker would stay with the company. But this variable was also strongly associated with race (Bogen and Rieke, 2018[20]; Yam and Skorburg, 2021[52]). In another example, a CV screening company found that the name "Jared" and playing high school lacrosse were strong signals of success. However, in and of themselves, these factors have no causal link to job performance, and yet they are strongly related to race (Gershgorn, 2018[161]).

---

[14] As pointed out by Susan Scott-Parker, founder of the Disability Ethical AI? Alliance (DEAI), one of the key challenges for people with disabilities in addition to the inevitable disability bias in the data, is that many AI recruitment tools use a standardised process. She argues that standardised recruitment processes are inherently discriminatory. Disability equality is only possible when disabled job seekers are provided with the reasonable adjustments, at every stage of a recruitment process, that enable them to compete on an equal basis.

[15] Ironically, not only are low-skill workers more likely to face difficulty in navigating such systems, but they are also more likely to be exposed to them when applying for a job. For example, Ajunwa (2016[154]) argues that the low-wage and hourly workforce is more likely to have to face AI as an initial hurdle to clear to gain employment. Job applications for some retail jobs must now be submitted online, where they will first be sorted by automated hiring platforms powered by algorithms. This might be related to the fact that the volume of hiring in these low-skill occupations is higher, so the potential efficiency gains from using AI are greater. It could also be because performance in low-skilled occupations might be easier to measure, which means that AI tools used in recruitment might be more accurate in the case of low-skilled workers.

Bias can also arise through the use of an algorithm. Some AI tools learn from recruiter and jobseeker actions, and this can perpetuate bias (Cadwalladr, 2016[162]). For example, Chen et al. (2018[163]) found that some of the gender bias embedded in CV search engines like Indeed, Monster and CareerBuilder, was due to the algorithms adjusting the ranking of candidates based on the volume of clicks they received from recruiters. If recruiters are biased (e.g., prefer men for certain roles), they may generate more clicks on candidates with certain demographic traits (e.g., men) and be shown more men by the algorithm in the future. Similarly, if a female candidate were more likely to click on junior positions because she doubts she is qualified enough for more senior positions, then the algorithm will be more likely to show her junior positions in the future. This may even be the case if she looks for positions at the right level, but other women look for more junior positions, and the algorithm puts her in the same demographic group as these other women (Bogen and Rieke, 2018[20]).[16]

Finally, bias can arise for unpredictable reasons, even when the AI is designed to be fair. Some examples come from online job advertising, which is often organised as auctions where advertisers bid for a target audience with certain desired demographic characteristics. Celis, Mehrotra and Vishnoi (2019[164]) give the following example of how targeted advertising could lead to discriminatory outcomes:

> *"Consider the setting in which there are two advertisers with similar bids/budgets, but one advertiser specifically targets women (which is allowed for certain types of ads, e.g., related to clothing), while the second advertiser does not target based on gender (e.g., because they are advertising a job). The first advertiser creates an imbalance on the platform by taking up ad slots for women and, as a consequence, the second advertiser ends up advertising to disproportionately fewer women and is inadvertently discriminatory."*

Lambrecht and Tucker (2019[152]) find something similar when they test an algorithm to deliver advertisements promoting Science, Technology, Engineering and Mathematics job opportunities. The advertisement was set up to be gender neutral in its delivery. However, in practice, fewer women saw the advertisement than men. The authors found that this was because younger women are a prized demographic and are more expensive to show advertisements to. An algorithm that simply optimises cost-effectiveness in the delivery of advertisements will end up being discriminatory because of crowding out.

Views on whether AI is more or less fair than humans are mixed. Although there is some evidence that, in practice, algorithms can be less biased and more objective than humans (Bornstein, 2018[48]; van Esch, Black and Ferolie, 2019[165]) and that HR professionals agree that the use of AI would result in more equitable hiring practices (Eightfold, 2021[166]), people's opinions are currently mixed as to whether algorithms are fairer than human beings or not (Box 12). Such perceptions of fairness matter for matching because, where applicants think algorithms are unfair, this can have important consequences for whether they decide to stay in the applicant pool or accept a job offer (Köchling and Wehner, 2020[167]).

Building entirely bias-free algorithms may not be possible, therefore the focus should be on reducing bias, and not eliminating it. A privacy director of a PrES interviewed as part of this project said:

> *"We have to accept that the world we live in is not without bias and that, for a long time to come, it will be impossible to create bias-free products. When we admit that, regulation can focus on making companies understand bias, how to audit for bias on an ongoing basis, and to implement measures to reduce bias as much as possible."*

This is a view echoed by the National Institute of Standards and Technology in the United States: "Bias is neither new nor unique to AI and it is not possible to achieve zero risk of bias in an AI system". The focus

[16] Note that the funding model of job platforms may encourage this type of behaviour, since they are often paid based on the number of clicks a job advertisement will get, and not on which candidate will ultimately be most successful in the job (Morrison & Foerster, n.d.[219])

of NIST is therefore “to develop methods for increasing assurance, governance and practice improvements for identifying, understanding, measuring, managing, and reducing bias” (Schwartz et al., n.d.[168]).

Box 12. The evidence is mixed as to whether AI is perceived as more or less fair than humans

Some studies have found that people consider AI to be fairer than humans. For example, participants in the study by Langer et al. (2020[106]) perceived highly automated interviews as more consistent than human beings. Bigman et al (2020[169]) showed that, in the recruitment context, people were less outraged by algorithmic than by human decision-making, and that this was because they ascribed less discriminatory intent to algorithms. Similarly, Tambe, Cappelli and Yakubovich (2019[67]) reported that people found decisions easier to accept when they were made by a machine, especially when they had negative consequences.

Other studies, however, conclude that people find algorithms to be less fair. Both Acikgoz et al. (2020[108]) and Bigman et al. (2020[108]) concluded that AI-based interviewing was viewed as less just than traditional human-based interviewing/selection, and Fast and Harmon (2020[170]) found that this was primarily because of perceptions of reductionism (see discussion on robustness above). Dineen, Noe and Wang (2004[171]) also found that, in a web-based applicant screening context, automated decision-making was perceived as less fair than human decision-making, and Langer, König and Papathanasiou (2019[103]) found that people reacted negatively to automated interviews, and more so the higher the stakes. Similarly, both Lee (2018[148]) and Newman, Fast and Harmon (2020[170]) concluded that algorithmic decisions were perceived as less fair than human ones.

A number of factors appear to impact people’s views of the fairness of algorithms. Wang, Harper and Zhu (2020[172]) suggest that there might be “outcome favourability” bias, i.e. that people rate the algorithm as more fair when the outcome is in their favour. Kaibel et al. (2019[173]) found that applicants who had experienced discrimination by humans perceived algorithms to be fairer, and there is also some evidence that perceptions of fairness depend on who the developer is. Wang, Harper and Zhu (2020[172]), for example, find that people’s perception of fairness of an algorithm is lower when it is built by an outsourced team, as opposed to when it is built within the organisation. This is one of the main reasons the French PES Pôle Emploi, has chosen to invest in AI skills and develop its own AI applications. In interviews held with Pôle Emploi as part of this project, they said that effective use of AI in the field can only be guaranteed if there is trust in the tools and, one way of obtaining that trust, is to develop the tools internally.

Finally, a few studies find no difference in fairness perceptions, including: Suen, Chen and Lu (2019[174]) for decision-making during video interviews; and Ötting and Maier (2018[175]) for perceptions of justice between human, robot and computer decision agents.

Human rights: Privacy

The use of AI in matching has raised questions around privacy. There is evidence, for example, that automated interviews lead to an increase in privacy concerns on the part of candidates (Langer, König and Papathanasiou, 2019[103]), which could be related to worries about surveillance as well as about who would access the data (Langer et al., 2020[106]). Moore (2020[97]) adds to this concerns around ownership of the data and around the use of data for purposes other than for which it was intended. This latter issue is a particular concern when it comes to data scraped from social media profiles (Hutchinson, 2022[176]). Technically, the use of such data is legal as long as users have agreed to the terms and conditions of the platform and/or have given the required permissions as part of their settings. Even so, individuals have concerns about the use of social media data for recruitment purposes. 50% of US adults said they would feel worse about computer programmes that evaluate job candidates if the programme used publicly available data like social media posts (Smith and Anderson, 2017[99]). A particularly worrying phenomenon is the use of social media data to infer protected characteristics. This is not only a privacy concern, but also a possible source of discrimination. For example, it has been found that Facebook likes could predict with a high degree of accuracy sensitive characteristics like gender, ethnicity, sexual orientation, religious and political views, etc. (Kosinski, Stillwell and Graepel, 2013[177]). Another study found that even just 100 Facebook likes could lead to the identification with high accuracy of a person's skin colour, sexual orientation, or political affiliation (Grassegger and Krogerus, 2018[178]). Big data has even been used to infer women's likelihood of becoming pregnant (Oswald et al., 2020[179]). Image and voice recognition techniques are similarly being used to infer information about applicants' sexual orientation, race, age, as well as physical attractiveness (Chamorro-Premuzic et al., 2016[21]; Dattner et al., 2019[180]).

Transparency and explainability

Transparency can be a challenge when it comes to the use of AI in labour market matching.[17] When AI is being used in recruitment, candidates are not always informed that this is the case, and some legislatures have already taken steps to address this (see Section 5). Moreover, even when candidates are informed, they are not always aware of what the AI is assessing. For example, Bodie et al. (2017[22]) look at AI-based games used in screening, and point out that it is not always clear what skills these tools are testing.

AI tools are often a black box and it can be hard to explain outcomes. AI algorithms are atheoretical. They look for patterns and correlations, but attach no importance as to whether these are meaningful in any way (Kim, 2017[46]). With increased analytical power and the growing size of datasets, the recommendations made by AI tools are becoming increasingly opaque (Giermindl et al., 2021[75]). Moreover, these tools adapt to and learn from each new data point, making explanations even more evasive (Faraj, Pachidi and Sayegh, 2018[159]). This also makes it difficult to spot errors in the system. As a result, it might be difficult for AI-based decisions to meet the transparency and explainability requirements of the General Data Protection Regulation (GDPR).

That being said, some authors have argued that transparency of algorithms may be a red herring. Zerilli et al. (2018[181]) argue that automated decision-making is being held to an unrealistically high standard, and that human decision-making processes are also black box decisions, so why expect so much more from automated decision-making? Moreover, Ananny and Crawford (2016[182]) believe that transparency alone will not achieve what is ultimately more desirable, i.e., accountable AI systems. This is an issue that will be explored in further detail in Section 5. Suffice it to say here that this is an important policy question, since it speaks to the efficacy of existing laws. If algorithms mean that recruiters can evade legal responsibility, then there may be a need to review that legislation.

[17] According to the OECD AI Principles (OECD, 2019[214]), transparency and explainability are two different, but related, concepts, which also can have different meanings. First, transparency can refer to disclosing when AI is being used. Transparency further means enabling people to understand how an AI system is developed, trained, operates, and is deployed, so that people acquiring or using these tools can make informed choices. Transparency also refers to the ability to provide meaningful information and clarity about what information is provided and why. Explainability means enabling people affected by the outcome of an AI system to understand how the decision was arrived at. This entails providing easy-to-understand information to people affected by an AI system's outcome that can enable those adversely affected to challenge the outcome, notably – and to the extent practicable – the factors and logic that led to an outcome.

5 Recent Policy Developments with an Impact on the Use of AI in Matching

This final section provides a brief overview of some of the key policy actions countries are undertaking to address the risks of AI discussed in Section 4 of the report. These actions include: promoting transparency and requiring consent where AI tools are being used; ensuring that there is a human in the loop and giving individuals a right to contest automated decisions; guaranteeing privacy; and fighting bias and discrimination.[18] The discussion centres around the application of these policy actions to the matching sphere. For a broader discussion of recent policy developments on AI and the labour market more broadly, the reader is referred to Salvi del Pero, Wyckoffi and Vourc'h (2022[1]).

Policy action on AI does not occur in a regulatory vacuum. While efforts are underway to regulate AI (e.g., the AI Act and the platform directive in the EU; as well as various state initiatives in the US), it is important to point out that, in most countries, the development and use of AI will already be regulated to some extent by existing legislation in the areas of: data protection, consumer protection, non-discrimination, and gender equality, amongst others. Misuraca and van Noordt (2020[183]) warn against a "gold rush to become a rule-maker in the field of AI governance" with documents that "for the most part – omit or override existing governance mechanisms and institutions, as if they were completely mismatched for 'the age of AI'". This is a situation to be avoided, and the first step in regulating AI should consist of looking at loopholes and ambiguities in existing legislation.

Regulation can come with a price, since there are compliance costs and it can reduce investments in AI. For example, it has been estimated that the implementation of GDPR had a pronounced short-term negative effect on venture investment (Jia, Jin and Wagman, 2019[184]). Similarly, some have argued that the European Unon's proposed AI Act could bring costs to companies and reduce AI investments (Mueller, 2021[185]). A managing director for workforce transformation at a large consulting company said:

> *"Regulation can be a big barrier. It increases the fear factor because it sends a signal that there are more challenges than organisations were aware of. Organisations, and particularly smaller ones, often cannot afford the time to delve into the regulation and to figure it out, so they have a tendency to wait and see what happens in the market."*

Uncoordinated regulation across regions could also be detrimental. A privacy director of a PrES interviewed as part of this project gave the example of the United States where some states now require Data Protection Impact Assessments, and he said there is a risk that different regulations across different states would place a big burden on smaller firms in particular.

However, regulation can also have a positive impact on the adoption of AI. Regulatory uncertainty can reduce investment in new technologies, as companies want to avoid potentially falling foul of the law, which could have both financial and reputational consequences (Li et al., 2021[59]). Users of AI need clear

[18] In addition to these interventions, the banning of certain AI tools may be considered. For example, the Council of Europe has called on European countries to impose a strict ban on facial analysis tools that purport to "detect personality traits, inner feelings, mental health or workers' engagement from face images" (Council of Europe, 2021[220]).

boundaries to operate—a point made repeatedly during interviews held with PrES as part of this project, all of whom were keenly watching regulatory developments in the United States, as well as the EU AI Act. In this sense, regulation could in fact promote the adoption of AI, since it is in the interest of users that the AI that is put on the market is trustworthy. Also, and in response to the concern that regulation might be costly for companies to comply with, it is worth remembering that the same companies are set to gain from these new technologies in terms of lower costs and higher efficiency, out of which any compliance costs with the regulation could be financed (Ajunwa, 2016[154]).

Transparency and consent

Regulators are increasingly insisting on transparency in the use of AI in recruitment and on the need to obtain consent. In the European Union, the GDPR covers the right to transparency on the use of automated decision-making and legal bases for lawfully processing personal data.[19] The proposed AI Act says that "when persons interact with an AI system or their emotions or characteristics are recognised through automated means, people must be informed of that circumstance."[20] Similarly, the EU's draft platform directive requires digital labour platforms to inform platform workers of the use (and key features) of automated monitoring and decision-making systems.[21] In the United States, some state-level initiatives are requiring recruiters to inform candidates when AI is being used in the recruitment process. In 2019, the State of Illinois introduced the Artificial Intelligence Video Interview Act (ILCS, 2019[186]) which requires employers to inform candidates prior to the interview that AI is being used, including how it works and what variables are under scrutiny, as well as to obtain written informed consent from the candidate. In 2020, the State of Maryland banned the use of facial recognition in job interviews, unless the candidate signs a waiver. New York City has also passed a bill on automated employment decision tools, which will require candidates and employees to be notified about the use of such tools for hiring or promotion, and about the job qualifications and characteristics used by the tool. Finally, the California Privacy Rights Act will come into force in 2023 and will require the state's Attorney General to promulgate regulations that govern individuals' opting out of automated decision-making, and that require businesses to provide "meaningful information about the logic involved in [automated] decision-making processes, as well as a description of the likely outcome of the process with respect to the consumer".

Transparency and consent may be particularly important for individuals who could be disadvantaged by such tools. People with disabilities are one such group. As an HR Innovation Strategist at a large consulting company said: "Many interfaces are not built for people with different disabilities, which makes it difficult for them to get jobs. We need to give people a right to know and a right to choose about when and how they interact with AI." In other words: AI could introduce accessibility barriers for

[19] GDPR does not require employers to seek consent for the use of automated decision tools. However, GDPR does include the right not to be subject to a decision based solely on automated processing if it risks producing legal effects concerning the data subject or similarly significant effects. As Sánchez-Monedero, Dencik and Edwards (2020[23]) argue, "given the importance of access to employment, automated hiring systems almost certainly make a decision which has legal or significant effect." In fact, in the United Kingdom, the Information Commissioner's Office (ICO), which regulates data protection, gives "e-recruiting practices without human intervention" as a canonical example for the application of article 22 of the GDPR. Because of this, Sánchez-Monedero, Dencik and Edwards (2020[23]) believe that the use of a fully automated hiring tool could be refused by an applicant, who could ask for a human decision instead.

[20] All citations from the EU AI Act refer to the original proposal published in April 2021.

[21] The information to be provided includes: the categories of actions monitored, supervised and evaluated (including by clients) and the main parameters that such systems take into account for automated decisions. The regulation specifies in what form and at which point in time this information is to be provided and that it should also be made available to labour authorities and platform workers' representatives upon request.

certain individuals and, as was discussed above, it could also introduce bias against them. As a result, individuals need to be informed about the use of such tools and be given a choice. Susan Scott-Parker, founder of the Disability Ethical AI? Alliance (DEAI), argues that people with disabilities need to be given the opportunity, early on in the recruitment process, to ask for reasonable adjustments, noting that in many jurisdictions employers are already legally obliged to provide such accommodations.

Providing individuals with information and seeking consent only makes sense if it offers them a true choice about whether such tools are used. GDPR specifies that consent must be "freely given, specific, informed and unambiguous." However, applicants hold less power than recruiting firms. As Sánchez-Monedero, Dencik and Edwards (2020[23]) and Zwitter (2014[187]) argue, it is difficult to give explicit consent in the context of hiring. Even if applicants are informed enough to consent to the process, they may not be able to opt out without being (or feeling that they might be) disadvantaged in the process. This is indeed the view of the European Data Protection Board (previously the Art 29 Working Party), who state that it is "problematic for employers to process personal data of current or future employees on the basis of consent as it is unlikely to be freely given" (European Data Protection Board, 2020[188]). As Moore (2020[97]), points out, collective bargaining and workers' voice might help to some extent to achieve meaningful consent—although this would apply more to existing workers than to applicants to new jobs.

Existing legislation might contain loopholes that enable employers to avoid seeking consent. For example, in the context of GDPR, automated individual decision-making, including profiling, which produces legal (or similar) effects concerning the individual data subject, is allowed only when it is predicated on consent, where specific national or European Union legislation permits such processing (and the processing is subject to appropriate safeguards), or where such processing is "necessary for entering into or performance of a contract between an organisation and the individual" (art 22 (1)(a))14. While it may be difficult to see why the use of an automatic hiring tool might ever be "necessary" for setting up an employment contract, employers might be able to argue that processing large numbers of applicants is only possible with the use of such tools, and therefore consent is not required (Sánchez-Monedero, Dencik and Edwards, 2020[23]).

Finally, even where consent has been obtained, this does not necessarily mean that it is "ethical" to use such tools (see also Section 4). This is an issue that arises, for example, in the context of using social media data for recruitment purposes. While, legally, social media content is public and can therefore be freely used, there is still a question of whether it is right to use such data for hiring purposes when, in practice, users had other purposes in mind when they consented to their data being made publicly available.

Human in the loop and the right to contest

Many existing or proposed AI regulations and principles restrict the use of automated decision-making (particularly in high-risk situations) and insist on having a "human in the loop". For example, the EU's General Data Protection Regulation (GDPR) includes the right not to be subject to a decision based solely on automated processing. The EU's proposed AI Act would require "human oversight throughout the AI systems' lifecycle", and the proposed directive on working conditions of people working through digital platforms expects human oversight and monitoring with respect to automated decision-making systems. In the United States, von Lewinski and de Barros Fritz (2022[189]) argue that Article 6 of the US constitution grants the right to a (human) jury. Similarly, the Algorithmic Accountability Act, if passed, could allow individuals to contest the use of algorithms.

Human oversight is meant to ensure that the decisions of an automated tool are checked, which could help attenuate risks. As Clyde (2021[190]) puts it, "The human-in-the-loop concept refers to inserting a human in between the machine and the outcome of its function in the world. Some say this hybrid decision-making model attempts to maintain human agency and accountability compared to the alternative automated-only systems." Joaquin Quinonero Candela from LinkedIn said, "AI should be used as a tool supporting recruiters' efforts to broaden and diversify their candidate pools. One of our key principles is that AI should not make hiring decisions. Humans should do that." And the Manager of Digital Ethics at a large consultancy company echoed this by saying, "AI can add value at all stages of the recruitment process, but always and primarily as an add-on to human judgment." There is evidence that job applicants are more positive about automated decisions when humans remain in control and are simply augmented by AI (Langer and Landers, 2021[100]) and companies may be of the same view. A Digital Transformation Leader at a large consulting company interviewed as part of this project said: "Companies want to keep the role of the recruiter in the middle. No organisation will totally relinquish control over hiring decisions to AI, they will always want to keep accountability and control."

In practice, there is a risk that human-in-the-loop arrangements could result in the mere rubber-stamping of automated decision-making. Sánchez-Monedero, Dencik and Edwards (2020[23]) argue that very few automated hiring and firing decisions seem to be taken without any human intervention at all, but that, in practice, this human intervention can mean very little. They cite the example of Amazon, who were accused of automatically firing over 10% of their employees when their productivity fell below a certain level. Amazon responded by saying that supervisors are always able to override the process—however this does not mean that it actually happens in practice. The risk of rubberstamping is real, given that people are subject to automation bias—although such bias can be mitigated through careful design (see Box 13)**.** Automation bias might be one reason why applicants perceive decisions to be fairer when recruiters only have the option to consult an automated system, as opposed to when they can only slightly change decisions that have already been made by an automated system (Newman, Fast and Harmon, 2020[170]). Similarly, Suen, Chen and Lu (2019[174]) find no negative reactions from candidates to algorithmic decision-making in personnel selection and argue this might be because algorithmic evaluation only served as a reference for the human decision-maker. Recruiters themselves have also been shown to be more satisfied with personnel selection decisions when they receive a ranking of applicants from an automated support system *after* they had processed the applicant information themselves. This had led to calls to flip the process on its head: instead of asking AI to surface the best candidates and having a human hiring manager rubber stamp its decisions, companies should use AI to audit their own recruiting practices (WEF, 2020[39]).

Box 13. Automation bias and how to address it

Humans display automation bias

"Automation bias"[22] is a term used to describe situations where "a human decision maker disregards or does not search for contradictory information in light of a computer-generated solution which is accepted as correct" (Cummings, 2004[191]).

In the context of personnel selection, researchers have found that when people are provided with a decision aid, their predictions are significantly more similar to (but not the same as) the predictions made by the aid, than when people are not provided with the decision aid (Jackson, 2016[192]). There is also evidence that people suffer from "position bias", i.e., they focus disproportionately on recommendations that appear at the top of a list than on those that appear further down (Joachims et al., 2005[193]). This matters because it may have an influence on how the AI tools and their output are best designed.

[22] Other, related terms include: automation-induced "complacency", "over-reliance" on automation, "automation dependence", and computer-induced "confirmation bias".

Automation bias diminishes the cognitive effort that individuals make to seek other information and it encourages them to make decisions too hastily after being prompted with an automated recommendation (Mosier et al., 1998[194]). Automation bias can result in either "commission errors" (when users follow an automated directive without taking into account other sources of information) or "omission errors" (when automated devices fail to detect or indicate problems and the user does not notice because they are not properly monitoring the system). For example, if a spell-checking programme incorrectly marks a word as misspelled and the user accepts the suggested alternative, then this would be a commission error. If, on the other hand, the spell-checking programme fails to notice a spelling error, then this would be an omission error.

Such errors can have various degrees of seriousness. In the case of a spelling recommendation, the consequences are obviously not severe. However, in the case of a pilot flying a plane or a doctor making a medical diagnosis, the consequences can be life-threatening. Similarly, in the employment context, automation bias could result in errors with significant consequences for the individuals concerned.

Errors resulting from automation bias need to be set against the reduced error rate that the automated recommendations are intended to achieve, so that what matters is the net change in errors as a result of the adoption of automation. If the automated tool is not perfect, then individuals in non-automated settings have been shown to out-perform their counterparts using the automated tool (Skitka, Mosier and Burdick, 1999[195]).

Automation bias can be mitigated through careful use and design

Individuals tend to over-rely on automated decision-making tools when their workload is heavy (Parasuraman and Riley, 1997[196]; Goddard, Roudsari and Wyatt, 2012[197]). Parasuraman and Manzey (2010[198]) find that automation complacency is more common under conditions of multiple-task load, when manual tasks compete with the automated task for the individual's attention. Lyell and Coiera (2017[199]) argue it is not so much multitasking, but rather the degree of cognitive load experienced in decision tasks that determines the degree of automation bias. Goddard, Roudsari and Wyatt (2012[197]) further identify task complexity and time constraint as factors influencing automation bias. Strategies to minimise automation bias should therefore focus on cognitive overload and making sure that workers are not pressured into accepting automated decisions.

A second issue that helps reduce automation bias, is to make individuals accountable for their decisions (Goddard, Roudsari and Wyatt, 2012[197]; Skitka, Mosier and Burdick, 2000[200]). There is, of course, a question about where this accountability should lie: with the direct user, or with those that have power (e.g., the executives of the company and/or the developers of the tools). This issue of accountability will be discussed further below in the report.

Finally, both the reliability and the design of the tools matter. Automation bias is more likely to occur where reliability and consistency of the tools are high (Parasuraman and Riley, 1997[196]; Alberdi et al., 2009[201]). Similarly, if confidence levels are attached to the recommendation, or if the tool provides "information" as opposed to a "recommendation", the automation bias tends to be lower (Goddard, Roudsari and Wyatt, 2012[197]). A HR Innovation Strategist at a large consulting company interviewed as part of this project said, "Some recruiters have algorithm bias, they trust whatever the tool tells them. This is partly down to the design of the tools which give numeric scores to candidates without margin of error, or any indication as to the precision of that score. When AI tools present a numeric score, it makes the assessment seem more precise than it really is." Kim and Duhachek (2020[202]) find that individuals are more likely to be persuaded by automated decisions when the system emphasises more concrete suggestions for how to perform a task. Automation bias is also found to be more likely, the higher the level of automation (Meyer, Feinshreiber and Parmet, 2003[203]; Cummings, 2004[191]).

By contrast, even though some studies have found that training (including exposing participants to rare automation failures) can reduce complacency (Bahner, Hüper and Manzey, 2008[204]) and, in particular, making individuals aware of the logic employed by the tool (Goddard, Roudsari and Wyatt, 2012[197]), others have shown that training is only partially effective (Mosier et al., 2009[205]). Making people work in teams (as opposed to on their own) also has no significant effect on automation bias (Skitka et al., 2000[206]).

One way of holding employers accountable for decisions taken by AI would be to give applicants a right to an explanation, as well as a right to contest an automated recruitment decision. In Europe, GDPR provides this right, and the proposed platform directive would further strengthen it by giving platform workers the right to obtain an explanation from the digital labour platform for a decision taken or supported by automated systems that significantly affects their working conditions. Where the explanation obtained is not satisfactory or where platform workers consider their rights infringed, they also would have the right to request the digital labour platform to review the decision.

Data protection and privacy

In the EU, the GDPR addresses many concerns around data protection and privacy that the use of AI in matching might raise. The GDPR lays down rules for the protections concerning the processing of personal data and expects the data controller to implement suitable measures to safeguard the data subject's rights and freedoms.[23] Such legislation does not yet exist in all OECD countries. In the United States, for example, there is no federal equivalent to the GDPR. Instead, there are different regulations and laws set by individual states and industry-based regulatory bodies. Some of these come close to GDPR standards (e.g. California, Colorado, Virginia) (IAPP, n.d.[207]), but others do not.[24] Overall, privacy laws in the United States tend to be relative loose compared to those in Europe (Bodie et al., 2017[22]). Von Lewinski and de Barros Fritz (2022[189]) point out that the Algorithmic Accountability Act 2019 would require certain companies in the US to conduct an automated decision-making system impact assessment to evaluate "the extent to which an information system protects the privacy and security of personal information the system processes". Canada is working on addressing AI systems in its privacy and data protection laws through Bill C-11 (2020) in Canada, which would require companies, upon request, to provide individuals with an explanation of decisions made by AI systems (Salvi del Pero, Wyckoffi and Vourc'h, 2022[1]).

In some countries, legislation on the use of social media data is still limited. For example, it is only in 2012 that lawmakers in the United States started introducing legislation to prevent employers from requesting passwords to personal social media accounts from applicants and worker (NCSL, n.d.[208]). Social media platforms themselves have taken some action (partly in response to the Cambridge Analytica scandal[25] as well as laws like the GDPR) (Chen, 2021[209]). Some social media companies are barring background check vendors from accessing their users' data (Bogen and Rieke, 2018[20]). In addition, some

[23] The proposed platform directive in the EU would strengthen restrictions around the use of personal data for platform workers. In particular, the regulation provides that digital labour platforms must not process any personal data concerning platform workers that are not intrinsically connected to and strictly necessary for the performance of their contract. This includes data on private conversations, on the health, psychological or emotional state of the platform worker and any data while the platform worker is not offering or performing platform work.

[24] In addition, the Illinois Artificial Intelligence Video Interview Act expects employers to limit the sharing of video interviews and destroy videos and copies of videos within 30 days up the applicant's request.

[25] In the 2010s, personal data belonging to millions of Facebook users was collected without their consent by British consulting firm Cambridge Analytica, predominantly to be used for political advertising.

jurisdictions give individuals “the right to be forgotten” so that any digital traces can be erased after some period of time and, in the future, it is likely that it will be more difficult for platforms to sell their data, or for employers to use them in recruitment.

Bias and discrimination

Approaches to tackle bias and discrimination resulting from the use of AI differ across countries and regions. They have included a mixture of: (i) anti-discrimination law; (ii) data protection legislation; (iii) consumer protection / product description legislation; and (iv) continued monitoring of algorithms throughout their lifetime (e.g., through audits). In fighting bias and discrimination in algorithms, it is important to remember that “AI operates in a larger social context” (Schwartz et al., n.d.[168]) and that a broad-based approach will be needed. In addition, it is important to see AI not just as the problem, but also as part of the solution.

Anti-discrimination legislation

While, in practice, the jury is still out on whether existing anti-discrimination law in the United States would be able to handle cases of algorithmic discrimination, there are some concerns. (Bornstein, 2018[48]). Title VII of the Civil Rights Act of 1964 forbids employers in the United States from discriminating on the basis of race, colour, religion, sex, and national origin.[26] It is conventionally understood to prohibit two kinds of discrimination: disparate treatment, and disparate impact. The former refers to intentional, or overt, discrimination, while the latter covers situations that are facially neutral, but nonetheless harm some workers more than others.

Challenging algorithmic discrimination as disparate treatment might be difficult under existing law (Bornstein, 2018[48])**.** This is because algorithms are “facially neutral” – i.e., they do not use protected characteristics to base decisions on. As a result, it cannot be argued that, if discrimination occurs, it is because an individual was intentionally treated differently based on a protected class. The “intention” to discriminate is important here because, even if apparently neutral data ends up being a proxy for a protected characteristic and results in discrimination, some have argued this can only be challenged as a case of disparate treatment if there was intention to do so—which will be very difficult to prove in practice (Kim, 2017[46]). Barocas and Selbst (2016[140]) have argued that a case of disparate treatment could be brought only if an employer wanted to hire on the basis of protected class membership and manipulated the algorithm to get the desired result. That being said, some commentators like Bornstein (2018[48]) believe that if AI is trained on data that reflects past discrimination and the AI reproduces such discrimination, then an employer who has intentionally chosen to feed “garbage in” to the model, may be litigable for disparate treatment under Title VII. Even Bornstein (2018[48]), however, acknowledges that, in practice, it might very difficult to bring such a case, since “Plaintiffs will likely need access to complex algorithms that may be inaccessible or from which individual factors may not be parsed, particularly if they involve “black box” machine learning”.

It might be more feasible to bring a case of disparate impact when there is algorithmic discrimination, but even here there may be some challenges. In particular, an employer’s use of algorithmic decision making could be excused and justified as “business necessity” when its outcomes are predictive of future employment outcomes (Bornstein, 2018[48]; Barocas and Selbst, 2016[140]). However, Kim (2017[46]) points out that, in the case of workforce analytics, the AI by definition relies on variables that are correlated in some sense with the job. So, “to ask whether the model is ‘job-related’ in the sense of ‘statistically correlated’ is tautological” (Kim, 2017[46]). Furthermore, Kim (2017[46]) points out that the text

[26] Some cities and states have expanded protections to characteristics not explicitly covered by Title VII, like gender identity, sexual orientation, citizenship status, and political affiliation (Bogen and Rieke, 2018[20]).

of Title VII forbids what she calls “classification bias” – i.e. the use of classification schemes that have the effect of exacerbating inequality or disadvantage along lines of race, sex, or other protected characteristics. As a result, Kim (2017[46]) believes that, “because the employer’s justification for using an algorithm amounts to a claim that it actually predicts something relevant to the job”, “the employer should carry the burden of demonstrating that statistical bias does not plague the underlying model. In other words, the employer should have to defend the accuracy of the correlations by showing that no problems exist with the data or model construction that are biasing the results.

AI developers in the United States are already taking steps to de-bias their models and reduce the risk of employers being exposed to disparate impact litigation. For instance, an AI developer and business strategist at an online job platform interviewed as part of this project mentioned two approaches for reducing bias in their recommender tool. The first approach, known as “post-processing”, involves setting quotas for the share of candidates recommended from each population sub-group. A common calibration for such tools ensures compliance with the 4/5 rule. This rule states that if the selection rate for one protected group is less than 4/5 of that of the group with the highest selection rate, the employer may be at risk of discrimination (Raghavan et al., 2019[54]).[27] The second approach involves feeding the algorithm artificial profiles to train it and boost the chances of under-represented groups being selected.

One challenge that developers face in ensuring that the outputs of their models are not biased, is that it is not clear whether they can use protected characteristics to calibrate their models’ results. This is because doing so might, itself, be a case of disparate treatment (an issue discussed further below in the section on algorithmic audits). According to Raghavan et al. (2019[54]), one approach taken by developers to get around this, is to take into account protected attributes when building the models, but ultimately produce models that do not take protected attributes as input. At an IZA workshop held in September 2022, Raghavan added that, in an attempt to sell their AI products, some vendors are taking on legal liability should a case of discrimination be brought against the employer.

Data protection legislation

Data protection laws might be as important in uncovering and combating bias in hiring tools as anti-discrimination laws are. In practice, it is difficult to spot discriminatory patterns in the decisions taken by AI. This is particularly the case in advertising/targeting of vacancies (individuals are not even aware that the vacancy was not shown to them). As Bogen and Rieke (2018[20]) put it, “While new hiring tools rarely make affirmative hiring decisions, they often automate rejections. Much of this activity happens early in the hiring process, when job opportunities are automatically surfaced to some people and withheld from others, or when candidates are deemed by a predictive system not to meet the minimum or desired qualifications needed to move further in the application process.”[28] This makes enforcement of anti-discrimination law difficult in practice. Part of the difficulty in spotting discriminatory patterns in the outputs of such tools lies in their complexity and opacity (Bogen and Rieke, 2018[20]). This is why regulatory efforts to boost transparency and explainability may also help in addressing bias and discrimination.[29] In Europe, the

[27] According to Manish Raghavan from MIT, one challenge with these approaches is that certain attributes of candidates are not collected, so it is impossible to know whether the 4/5 rule is being met. Another challenge is that, in some cases (e.g., certain disabilities), there simply would not be enough data to test for compliance with the 4/5 rule, because the number of applicants with that attribute would be too low.

[28] One example of this, is Facebook’s “Lookalike Audiences” feature, which allows recruiters to provide a list of their existing workers for Facebook to target a vacancy on Facebook users who are demographically similar (Ajunwa, 2016[154]). This feature has been used in the past to exclude older workers from seeing certain job ads (Angwin, Scheiber and Tobin, 2017[215]), or workers of a certain sex (Tobin and Merrill, 2018[216]).

[29] That being said, some have argued that, compared with human bias, algorithmic bias is easier to detect and, especially, to remove (Florentine, 2016[217]; Polli, 2019[218]).

GDPR gives individuals the right to transparency when it comes to their data, as well as the right to an explanation and to contest decisions —all of which would put pressure on employers to make sure that the hiring tools they use are not biased.[30] In addition, Sánchez-Monedero, Dencik and Edwards (2020[23]) argue that any machine learning system is likely to be regarded as "high risk" processing and will therefore require a Data Protection Impact Assessment, which should show, amongst others, that the risk of bias and discrimination has been mitigated.

Consumer protection / product description legislation

Consumer protection legislation might help ensure that AI tools that are put on the market are bias-free. In Europe, in addition to GDPR, the proposed AI Act would place responsibility on developers prior to putting an AI product on the market. In particular, the proposed legislation identifies AI systems used in employment as "high risk" and argues that they "may perpetuate historical patterns of discrimination". The Act would apply requirements to such systems as regards "the quality of data sets used, technical documentation and record-keeping, transparency and the provision of information to users, human oversight, and robustness, accuracy and cybersecurity". The emphasis on data quality is motivated by a desire to ensure that "the high-risk AI system […] does not become the source of discrimination prohibited by Union law". Box 14 provides more detail on the proposed AI Act.

Another way of putting responsibility on the vendor, is by regulating the descriptions they provide of AI products when they are sold. At the moment, vendors of AI technologies rarely disclose information about whether they validate their models, what validation methodologies they use, the validation data used, or how the validation procedures might be tailored to a particular client (Raghavan et al., 2019[54]). In some cases, vendors have even engaged in unfair or deceptive practices when selling their products, notably through providing inaccurate descriptions and/or a general lack of transparency about what their algorithms are doing (Federal Trade Commission, 2021[210]). All this can lead to unrealistic expectations on the part of the buyer about what AI can, and cannot, do. One example of deceptive practices emerged when, following a complaint to the FTC by the Electronic Privacy Information Center, HireVue announced that it would stop analysing facial expressions in videos to assess job candidates (Kahn, 2021[211]). The Center's complaint had called such practices deceptive alleging that, while marketed as being more objective than decisions made by human resource managers, HireVue's AI systems' decisions were in fact more likely to be biased. In 2021, the FTC approved new compulsory process resolutions in eight key enforcement areas, including bias in algorithms and biometrics, to enable more aggressive investigations of conduct and swifter action against companies in the US engaging in any conduct addressed by the resolutions. In the EU, the proposed AI Act would expect vendors to provide "technical documentation and record-keeping", as well as more "transparency and the provision of information to users" (Box 14).[31]

[30] In practice, however, it may be difficult for individuals to challenge AI decisions, especially in cases where they are not aware that AI is being used.

[31] While regulation in this area is critical, buyers themselves could also take steps to ensure that the tools they buy align with their values. The Manager of Digital Ethics at a large consultancy company said, "The question is whether the values of the vendor align with those of the firm purchasing the AI services. Firms could do an ethics impact assessment, which doesn't have to be very complex or onerous. A procurement manager could do it in a day."

Box 14. The proposed EU AI Act

One of the most significant regulatory developments in response to AI is the EU's Artificial Intelligence Act, the proposal for which was published in April 2021. It will still take years before any such legislation takes effect and the contents of the proposal are still subject to change. However, in its original form, the AI Act's objective is a set of harmonised rules for the development, placement on the market, and the use of AI systems in the European Union. It aims to ensure that AI systems placed on the market and used in the EU are safe and respect fundamental rights, such as the principle of equal treatment. The AI Act aims to address inherent challenges such as bias and lack of accountability, helping to tackle the risk of discrimination.

The AI Act follows a risk-based approach. It distinguishes between uses of AI that create: (i) an unacceptable risk; (ii) a high risk; and (iii) low or minimal risk. Unacceptable risk includes certain particularly harmful AI practices, such as: subliminal, manipulative or exploitative techniques causing harm; real-time, remote biometric identification systems used in public spaces for law enforcement; and all forms of social scoring.

The high-risk category is the most relevant one from a matching perspective, since it includes all AI systems used in employment, workers management and access to self-employment, and those systems may appreciably impact future career prospects and livelihoods of these persons. In particular, it includes: AI systems intended to be used for recruitment or selection of natural persons, notably for advertising vacancies, screening or filtering applications, evaluating candidates in the course of interviews or tests.

According to the proposal, such AI systems will have to comply with a set of mandatory requirements and follow conformity assessment procedures before they can be placed on the EU market. This includes requirements regarding the quality of data sets used, technical documentation and record-keeping, transparency and the provision of information to users, human oversight, and robustness and accuracy.

Continuous monitoring and audits

The idea of audits is gaining traction in the policy sphere as a tool to prevent the use of biased algorithms (in addition to other objectives that audits might achieve). Since some AI continuously evolves as it interacts and learns from real-world data, traditional product quality assurance (which focuses on product quality at the point of introduction to the market) would not be sufficient to guarantee the continued reliability of AI tools. This is why the idea of carrying out regular "audits" on AI tools has been floated. In some ways this would be similar to doing annual safety checks on gas boilers or cars, to make sure that the AI tool still does what it was intended to do. Audits could also be seen as an additional way of promoting transparency (Kim, 2017[212]). The EU's proposed AI Act discusses the harmonisation of the way in which ex-post controls are conducted once AI systems have been placed on the market. New York City was the first to introduce legislation which will require "automated employment decision tools" used in hiring and promotion to be audited (Box 15).

Box 15. Audits of automated employment decision tools in New York City

From January 2023 onwards, New York City will require “automated employment decision tools” used in hiring and promotion to be audited on an annual basis. Such audits will consist of an “impartial evaluation” conducted by an “independent auditor” that includes, at a minimum, an analysis of whether the automated employment decision tool has resulted in a disparate impact based on gender, race or national origin. The results these audits will need to be made publicly available on the website of the employer or employment agency. In the absence of such audits, an employer would not be able to use the tools and any employer failing to comply may be subject to a fine of up to USD 500 for a first violation and then penalised by fines of between USD 500 and USD 1 500 daily for each subsequent violation.

The effectiveness of audits is likely to depend on their design. A number of important questions need answering, including: (i) who is responsible for the audits (the developer or user); (ii) who carries out the audit, what professional accreditation do they require, and what is their degree of independence vis-à-vis the user and developer; (iii) what access does the auditor get to proprietary data/code; (iv) what standards are set for the audit (and what degree of flexibility do the auditors/auditees have in setting those standards); (v) the transparency of the process; and (vi) what are the consequences if no audit is carried out, or if the audit uncovers bias? In addition, it is important that responsibility/accountability is clearly allocated and that non-compliance is treated seriously. For example, Ajunwa (2016[154]) has argued that an employer’s failure to audit (and correct) its automated hiring tool should serve as prima facie evidence of discriminatory intent. The issue of accountability will be explored further in the next sub-section.

An outstanding question is whether taking action based on audit results might fall foul of anti-discrimination legislation. Kroll et al. (2017[213]) have argued that the very auditing of AI for discriminatory outcomes and making changes based on these findings would open an employer up to disparate treatment liability, since the employer would be tweaking the selection procedure based on a protected characteristic. This fear is based on a particular interpretation of the Supreme Court’s decision in Ricci v. De Stefano. In this case, the New Haven Fire Department had introduced a test for promotion. However, after finding that none of the Black firefighters who took the test scored high enough to be considered for promotion, New Haven officials invalidated the test results and the non-Black firefighters who passed the test were denied promotion. These non-Black firefighters subsequently sued the city and argued that, by discarding the test results, they had been discriminated against on the basis of race. The Supreme Court agreed that New Haven’s decision to ignore the test results was a case of disparate treatment. Subsequently, this decision been interpreted by some to mean that using protected characteristics to revise an algorithm and remove bias could also be a case of disparate treatment. However, as Kim (2017[212]) points out, one crucial difference is that, in Ricci v. De Stefano, the decision to drop the test was made *after* it had been fielded, whereas any changes to an algorithm as a result of an audit would be done *before* using the algorithm in recruitment. In fact, Kim (2017[212]) goes further and argues that some readings of Title VII of the Civil Rights Act have suggested that the law would in fact encourage (rather than ban) the use of protective characteristics to assess the risks that a model produces biased outcomes. In the EU, the proposed AI Act also appears to open the door for the use of protected characteristics to monitor, detect and correct bias in high-risk AI systems. The original draft text of the AI Act suggested that, “In order to protect the right of others from the discrimination that might result from the bias in AI systems, the providers should be able to process also special categories of personal data, as a matter of substantial public interest, in order to ensure the bias monitoring, detection and correction in relation to high-risk AI systems.”

The question of accountability

Underneath many of the issues discussed above, lurks the question of who would be accountable in the case something went wrong with the algorithm: the developer of the tool, or the user/recruiter.[32] The complication is that AI is not like traditional goods in the sense that it is not sufficient to test its quality or safety before it is put on the market. Many tools interact with and learn from new data, and so they change as they are used. This is a particular challenge for bias and discrimination, because bias could emerge in an AI tool either because of the way it was designed, or because of the way it is used. Although some developers, in an attempt to sell their tools, guarantee that they will take on responsibility if something goes wrong (see discussion of bias above), others are quick to absolve themselves. For example, HireVue's service agreement states that "Buyer understands and acknowledges that HireVue is solely a technology platform provider and does not participate in the interview, selection, or hiring of candidates. Accordingly, it is Buyer's sole responsibility to comply with all laws applicable to Buyer's use of the Service, including without limitation all applicable employment and hiring laws and regulations."

Approaches to accountability vary between Europe and the United States. The proposed EU AI Act would place much of the responsibility on the developer. Even when the product has been brought onto the market, "all providers should have a post-market monitoring system in place" as well as a system "to report to the relevant authorities any serious incidents or any breaches to national and Union law protecting fundamental rights resulting from the use of their AI systems". This is considered "key to ensure that the possible risks emerging from AI systems which continue to 'learn' after being placed on the market or put into service can be more efficiently and timely addressed". By contrast, the proposed EU Act places only limited, and very vague, responsibility on the user/employer: "Users should in particular use high-risk AI systems in accordance with the instructions of use and certain other obligations should be provided for with regard to monitoring of the functioning of the AI systems and with regard to record-keeping, as appropriate." More specifically, the AI Act would expect users of high-risk systems to use such systems in accordance with the instructions of use accompanying the systems. They would also need to ensure that input data is relevant in view of the intended purpose of the high-risk AI system, to monitor the operation of the high-risk AI system on the basis of the instructions of use, and to inform the provider or distributor of any serious incident or malfunctioning. The proposed AI Act would also place developer responsibility on users if they "modify the intended purpose of a high-risk AI system already placed on the market or put into service; [or] they make a substantial modification to the high-risk AI system". While the would AI Act place relatively few responsibilities on the user, it should be seen against the backdrop of GDPR, which already places some responsibility on the user (see discussion on data protection and privacy as well as keeping a human in the loop). In the United States, one of the main points of discussion is whether an employer would be responsible under anti-discrimination law if the decisions made by an AI tool were found to be discriminatory (see discussion on discrimination). In those discussions, there is little reference to the responsibility of the developer. Some developers themselves have been quick to shift the responsibility onto the employer[33] and the New York City Bias Law seems to place the burden on the employer (since they will be the ones paying fines if they are not in compliance with the law). At the same time, in the United States, there is increasing pressure on vendors to provide accurate descriptions of their AI products and provide transparency about what the algorithms are doing (see discussion on product description law).

[32] The proposed EU AI Act also discusses the responsibilities of importers, distributors, and authorised representatives.

[33] For example, in its service agreement, the company HireVue states that the "Buyer understands and acknowledges that HireVue is solely a technology platform provider and does not participate in the interview, selection, or hiring of candidates. Accordingly, it is Buyer's sole responsibility to comply with all laws applicable to Buyer's use of the Service, including without limitation all applicable employment and hiring laws and regulations."

References

Accenture (2021), *Ask Larry: AI Virtual Assistant for Jobless Claims*, https://www.accenture.com/us-en/case-studies/public-service/twc-unemployment-ai (accessed on 22 December 2022). [27]

Accenture (2015), *Do More With Analytics. Do Better With Labour Market Interventions*, https://www.accenture.com/no-en/insight-analytics-better-labour-market-interventions (accessed on 7 March 2022). [72]

Acikgoz, Y. et al. (2020), "Justice perceptions of artificial intelligence in selection", *International Journal of Selection and Assessment*, Vol. 28/4, pp. 399-416, https://doi.org/10.1111/IJSA.12306. [108]

Ajunwa, I. (2016), "The Paradox of Automation as Anti-Bias Intervention", *Cardozo Law Review*, Vol. 41, pp. 1671-1742, https://doi.org/10.2139/SSRN.2746078. [154]

Alberdi, E. et al. (2009), *Why are people's decisions sometimes worse with computer support?*, https://doi.org/10.1007/978-3-642-04468-7_3. [201]

Allegis Group (2017), *AI's Impact on Jobs, Talent, & HR*, https://www.allegisgroup.com/insights/blog/2017/august/ai-impact-on-jobs-talent-hr (accessed on 3 May 2022). [86]

Allred, C. (2019), *Applying a Metaheuristic Algorithm to a Multi-Objective Optimisation Porblem Within Personnel Psychology*, North Carolina State University, Raleigh, North Carolina, https://repository.lib.ncsu.edu/bitstream/handle/1840.20/37053/etd.pdf?sequence=1 (accessed on 18 October 2022). [58]

American Staffing Association (2016), *Personal Contacts Lead to Jobs*, https://americanstaffing.net/posts/2016/06/16/personal-contacts-lead-jobs/ (accessed on 10 February 2022). [101]

Ananny, M. and K. Crawford (2016), "Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability", *New Media and Society*, Vol. 20/3, pp. 973-989, https://doi.org/10.1177/1461444816676645. [182]

Ángeles Quiroga, M. et al. (2015), "Can we reliably measure the general factor of intelligence (g) through commercial video games? Yes, we can!", *Intelligence*, Vol. 53, pp. 1-7, https://doi.org/10.1016/J.INTELL.2015.08.004. [133]

Angrave, D. et al. (2016), "HR and analytics: why HR is set to fail the big data challenge", *Human Resource Management Journal*, Vol. 26/1, pp. 1-11, https://doi.org/10.1111/1748-8583.12090. [77]

Angwin, J., N. Scheiber and A. Tobin (2017), *Dozens of Companies Are Using Facebook to Exclude Older Workers From Job Ads*, ProPublica, https://www.propublica.org/article/facebook-ads-age-discrimination-targeting (accessed on 29 March 2022). [215]

Aranguiz, A. (2021), *Spain's platform workers win algorithm transparency*, Social Europe, https://socialeurope.eu/spains-platform-workers-win-algorithm-transparency (accessed on 17 August 2022). [9]

Awol Group (2019), *Artificial Intelligence in Employment Services - A Mapping*, https://tem.fi/documents/1410877/15020328/Artificial+intellingence+in+employment+services+-+A+mapping/24844ede-0570-c8da-4ed3-c91ec25b8e76/Artificial+intellingence+in+employment+services+-+A+mapping.pdf (accessed on 18 January 2022). [69]

Azucar, D., D. Marengo and M. Settanni (2018), "Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis", *Personality and Individual Differences*, Vol. 124, pp. 150-159, https://doi.org/10.1016/j.paid.2017.12.018. [126]

Bacchi, U. and A. Asher-Schapiro (2020), *The gig workers taking legal action to regain control of their data*, Reuters, https://www.reuters.com/article/global-workers-data-idUSKBN28Q0OY (accessed on 17 August 2022). [10]

Bader, B. (2019), *How do HR professionals feel about artificial intelligence?*, People Management, https://www.peoplemanagement.co.uk/article/1743841/how-does-hr-feel-about-artificial-intelligence (accessed on 3 May 2022). [66]

Baert, S. (2017), "Hiring Discrimination: An Overview of (Almost) All Correspondence Experiments Since 2005", *IZA Discussion Paper Series*, https://docs.iza.org/dp10738.pdf (accessed on 18 August 2022). [47]

Bahner, J., A. Hüper and D. Manzey (2008), "Misuse of automated decision aids: Complacency, automation bias and the impact of training experience", *International Journal of Human Computer Studies*, Vol. 66/9, pp. 688-699, https://doi.org/10.1016/J.IJHCS.2008.06.001. [204]

Barocas, S. and A. Selbst (2016), "Big Data's Disparate Impact", *California Law Review*, Vol. 104, pp. 671-732, https://29qish1lqx5q2k5d7b491joo-wpengine.netdna-ssl.com/wp-content/uploads/2016/06/2Barocas-Selbst.pdf (accessed on 21 February 2022). [140]

Barrett, L. et al. (2019), "Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements", *Psychological Science in the Public Interest*, Vol. 20/1, pp. 1-68, https://doi.org/10.1177/1529100619832930. [120]

Bartsch, L. (2020), *Survey: AI in Recruiting & How Companies feel about it*, Viasto, https://www.viasto.com/en/blog/survey-artificial-intelligence-in-hr/ (accessed on 3 May 2022). [79]

Benfield, S. (2017), *How Artificial Intelligence Is Improving Talent Acquisition*, HR.com, https://www.hr.com/en/magazines/all_articles/how-artificial-intelligence-is-improving-talent-ac_j0c8jqu4.html (accessed on 19 January 2022). [33]

Biel, J., L. Teijeiro-Mosquera and D. Gatica-Perez (2012), *FaceTube: Predicting personality from facial expressions of emotion in online conversational video*, https://doi.org/10.1145/2388676.2388689. [110]

Bigman, Y. et al. (2020), "Algorithmic Discrimination Causes Less Moral Outrage than Human Discrimination", https://doi.org/10.31234/OSF.IO/M3NRP. [169]

Black, J. and P. van Esch (2020), “AI-enabled recruiting: What is it and how should a manager use it?”, *Business Horizons*, Vol. 63/2, pp. 215-226, https://doi.org/10.1016/J.BUSHOR.2019.12.001. [19]

Blanchard, O. et al. (1989), “The Beveridge Curve”, *Brookings Papers on Economic Activity*, Vol. 1989/1, p. 76, https://doi.org/10.2307/2534495. [3]

Bodie, M. et al. (2017), “The Law and Policy of People Analytics”, *University of Colorado Law Review*, Vol. 88, pp. 961-1042, http://www.wired.com/2014/10/faster-higher-stronger/ (accessed on 18 February 2022). [22]

Bogen, M. and A. Rieke (2018), *Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias*, Upturn, https://www.upturn.org/static/reports/2018/hiring-algorithms/files/Upturn%20--%20Help%20Wanted%20-%20An%20Exploration%20of%20Hiring%20Algorithms,%20Equity%20and%20Bias.pdf (accessed on 14 January 2022). [20]

Bornstein, S. (2018), “Antidiscriminatory Algorithms”, *Alabama Law Review*, Vol. 70/2, pp. 519-572, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3307893 (accessed on 14 April 2022). [48]

Boyd, R. et al. (2015), *Values in Words: Using Language to Evaluate and Understand Personal Values*, https://ojs.aaai.org/index.php/ICWSM/article/view/14589 (accessed on 28 January 2022). [125]

Burke, L. (2019), *As AI-assessed job interviewing grows, colleges try to prepare students*, Inside Higher Ed, https://www.insidehighered.com/news/2019/11/04/ai-assessed-job-interviewing-grows-colleges-try-prepare-students (accessed on 2 August 2022). [145]

Cadwalladr, C. (2016), “Google, democracy and the truth about internet search”, *The Guardian*, https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook (accessed on 15 April 2022). [162]

Cappelli, P. (2017), *There’s No Such Thing as Big Data in HR*, Harvard Business Review, https://hbr.org/2017/06/theres-no-such-thing-as-big-data-in-hr (accessed on 9 March 2022). [81]

CareerBuilder (2017), *More Than Half of HR Managers Say Artificial Intelligence Will Become a Regular Part of HR in Next 5 Years*, https://press.careerbuilder.com/2017-05-18-More-Than-Half-of-HR-Managers-Say-Artificial-Intelligence-Will-Become-a-Regular-Part-of-HR-in-Next-5-Years (accessed on 3 May 2022). [65]

Cecchi-Dimeglio, P. (2017), *How Gender Bias Corrupts Performance Reviews, and What to Do About It*, Harvard Business Review, https://hbr.org/2017/04/how-gender-bias-corrupts-performance-reviews-and-what-to-do-about-it (accessed on 15 April 2022). [137]

Celis, E., A. Mehrotra and N. Vishnoi (2019), *Toward Controlling Discrimination in Online Ad Auctions*, http://proceedings.mlr.press/v97/mehrotra19a/mehrotra19a.pdf (accessed on 29 September 2022). [164]

Center for Digital Government, IBM and NASCIO (2021), *AI Meets the Moment: How a Pandemic has Shaped AI Adoption in State Government and What it Means for the Future*, https://www.nascio.org/wp-content/uploads/2021/10/NASCIO-CDG-IBM-AI-Meets-the-Moment-2021.pdf (accessed on 22 December 2022). [28]

Chamorro-Premuzic, T. (2019), *Will AI Reduce Gender Bias in Hiring?*, Harvard Business Review, https://hbr.org/2019/06/will-ai-reduce-gender-bias-in-hiring (accessed on 6 May 2022). [50]

Chamorro-Premuzic, T. et al. (2017), "The datafication of talent: how technology is advancing the science of human potential at work", *Current Opinion in Behavioral Sciences*, Vol. 18, pp. 13-16, https://doi.org/10.1016/J.COBEHA.2017.04.007. [135]

Chamorro-Premuzic, T. et al. (2016), "New Talent Signals: Shiny New Objects or a Brave New World?", *Industrial and Organizational Psychology*, Vol. 9/3, pp. 621-640, https://doi.org/10.1017/IOP.2016.6. [21]

Chen, B. (2021), *The Battle for Digital Privacy Is Reshaping the Internet*, The New York Times, https://www.nytimes.com/2021/09/16/technology/digital-privacy.html (accessed on 22 April 2022). [209]

Chen, L. et al. (2018), "Investigating the Impact of Gender on Rank in Resume Search Engines", *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* 651, pp. 1-14, https://doi.org/10.1145/3173574. [163]

Clyde, A. (2021), *Human-in-the-Loop Systems Are No Panacea for AI Accountability*, Tech Policy Press, https://techpolicy.press/human-in-the-loop-systems-are-no-panacea-for-ai-accountability/ (accessed on 4 May 2022). [190]

Cockx, B., M. Lechner and J. Bollens (2020), "Priority to Unemployed Immigrants? A Causal Machine Learning Evaluation of Training in Belgium", University of Ghent, Ghent, http://www.researchgate.net/profile/Michael_Lechner. (accessed on 19 January 2022). [40]

Council of Europe (2021), *Consultative Committee of the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data: Convention 108: Guidelines on Facial Recognition*, https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090000 (accessed on 13 November 2022). [220]

Cowgill, B. (2020), "Bias and Productivity in Humans and Algorithms: Theory and Evidence from Résumé Screening", Columbia University, https://www.theguardian.com/world/2016/oct/27/angela-merkel- (accessed on 26 January 2022). [55]

Crawford, K. (2013), *Think Again: Big Data*, https://foreignpolicy.com/2013/05/10/think-again-big-data/ (accessed on 21 February 2022). [156]

Cummings, M. (2004), *Automation Bias in Intelligent Time Critical Decision Support Systems*, American Institute of Aeronautics and Astronautics, Reston, Virigina, https://doi.org/10.2514/6.2004-6313. [191]

Dahlbom, P. et al. (2020), "Big data and HR analytics in the digital era", *Baltic Journal of Management*, Vol. 15/1, pp. 120-138, https://doi.org/10.1108/BJM-11-2018-0393/FULL/PDF. [76]

Dastin, J. (2018), "Amazon scraps secret AI recruiting tool that showed bias against women", *Reuters*, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G (accessed on 21 February 2022). [150]

Datta, A., M. Tschantz and A. Datta (2015), *Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination*, https://doi.org/10.1515/popets-2015-0007. [151]

Dattner, B. et al. (2019), *The Legal and Ethical Implications of Using AI in Hiring*, https://hbr.org/2019/04/the-legal-and-ethical-implications-of-using-ai-in-hiring (accessed on 18 March 2022). [180]

DeGeurin, M. (2022), *Microsoft's Calling It Quits on Creepy Emotion Recognition Tech*, GIZMODO, https://gizmodo.com/microsoft-shutting-down-azure-facial-emotion-recognitio-1849090511?utm_campaign=Neural%20Newsletter&utm_medium=email&_hsmi=217380324&_hsenc=p2ANqtz-_r9Nejy-3XHOLOhZS_iOKF3kZ6zLCl7PPi7e2Sv4cUdSjZwgI4qcR1iHKgx3xduLFxgByFHc8auICaevCU0zdqo_VfJQ&utm_content=217380324&utm_source=hs_email (accessed on 18 August 2022). [121]

Desiere, S., K. Langenbucher and L. Struyven (2019), "Statistical profiling in public employment services: An international comparison", *OECD Social, Employment and Migration Working Papers*, No. 224, OECD Publishing, Paris, https://dx.doi.org/10.1787/b5e5f16e-en. [24]

Dietvorst, B., J. Simmons and C. Massey (2015), "Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err", *Journal of Experimental Psychology: General*, Vol. 144/1, pp. 114-126, https://doi.org/10.1037/xge0000033. [60]

Dineen, B., R. Noe and C. Wang (2004), "Perceived fairness of web-based applicant screening procedures: Weighing the rules of justice and the role of individual differences", *Human Resource Management*, Vol. 43/2-3, pp. 127-145, https://doi.org/10.1002/HRM.20011. [171]

Duarte, N., E. Llanso and A. Loup (2017), *Mixed Messages? The Limits of Automated Social Media Content Analysis*, Center for Democracy & Technology, https://cdt.org/insights/mixed-messages-the-limits-of-automated-social-media-content-analysis/ (accessed on 18 February 2022). [129]

Dupré, D. et al. (2020), "A performance comparison of eight commercially available automatic classifiers for facial affect recognition", *PLOS ONE*, Vol. 15/4, p. e0231968, https://doi.org/10.1371/JOURNAL.PONE.0231968. [112]

Eightfold (2021), *HR' s Future State Report 2021: The Impact of Artificial Intelligence on Talent Processes The Impact of Artificial Intelligence on Talent Processes*, https://eightfold.ai/wp-content/uploads/HR_Future_State_Report_2021.pdf (accessed on 3 May 2022). [166]

Endsley, M. (2017), "From Here to Autonomy: Lessons Learned from Human-Automation Research", *Human Factors*, Vol. 59/1, pp. 5-27, https://doi.org/10.1177/0018720816681350. [96]

European Commission, D. and W. Pieterson (2020), *Digital technologies and advanced analytics in PES*, https://op.europa.eu/en/publication-detail/-/publication/464d4180-e741-11ea-ad25-01aa75ed71a1/language-en (accessed on 18 January 2022). [68]

European Data Protection Board (2020), *Guidelines 05/2020 on consent under Regulation 2016/679*, https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf (accessed on 4 May 2022). [188]

Ewell, P. et al. (2016), “Good Person or Bad Character? Personality Predictors of Morality and Ethics in Avatar Selection for Video Game Play”, *Cyberpsychology, Behavior, and Social Networking*, Vol. 19/7, pp. 435-440, https://doi.org/10.1089/CYBER.2015.0207. [131]

Faraj, S., S. Pachidi and K. Sayegh (2018), “Working and organizing in the age of the learning algorithm”, *Information and Organization*, Vol. 28/1, pp. 62-70, https://doi.org/10.1016/J.INFOANDORG.2018.02.005. [159]

Federal Trade Commission (2021), *Aiming for truth, fairness, and equity in your company’s use of AI*, Federal Trade Commission, https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai (accessed on 24 November 2021). [210]

Feldstein, M. (1978), “The Private and Social Costs of UnemploymenT”, *The American Economic Review*, Vol. 68/2, pp. 155-158, https://www.jstor.org/stable/1816681#metadata_info_tab_contents (accessed on 9 August 2022). [5]

Fisher, S. and G. Howardson (2022), “Chapter 16: Fairness of artificial intelligence in human resources - held to a higher standard?”, in Strohmeier, S. (ed.), *Handbook of Research on Artificial Intelligence in Human Resource Management*, Elgar, https://www.elgaronline.com/configurable/content/edcoll$002f9781839107528$002f9781839107528.00028.xml?t:ac=edcoll%24002f9781839107528%24002f9781839107528.00028.xml (accessed on 10 June 2022). [51]

Florentine, S. (2016), *How artificial intelligence can eliminate bias in hiring*, CIO, https://www.cio.com/article/236913/how-artificial-intelligence-can-eliminate-bias-in-hiring.html (accessed on 6 May 2022). [217]

Foroughi, C. et al. (2016), “Can we create a measure of fluid intelligence using Puzzle Creator within Portal 2?”, *Intelligence*, Vol. 56, pp. 58-64, https://doi.org/10.1016/J.INTELL.2016.02.011. [132]

Fraij, J. and V. Lászlo (2021), “A literature Review: Artificial Intelligence Impact on the Recruitment Process”, *International Journal of Engineering and Management Sciences*, Vol. 6/1, pp. 108-119, https://ojs.lib.unideb.hu/IJEMS/article/view/8288/8452 (accessed on 3 May 2022). [84]

Freire, M. and L. de Castro (2021), “e-Recruitment recommender systems: a systematic review”, *Knowledge and Information Systems*, Vol. 63/1, pp. 1-20, https://doi.org/10.1007/S10115-020-01522-8/TABLES/7. [8]

Frith, U. (2015), *Unconscious bias*, The Royal Society, https://royalsociety.org/-/media/policy/Publications/2015/unconscious-bias-briefing-2015.pdf (accessed on 14 June 2022). [45]

Fritts, M. and F. Cabrera (2021), “AI recruitment algorithms and the dehumanization problem”, *Ethics and Information Technology 2021 23:4*, Vol. 23/4, pp. 791-801, https://doi.org/10.1007/S10676-021-09615-W. [147]

Fruchterman, J. and J. Mellea (2018), *Expanding Employment Success for People with Disabilities*, Benetech, https://benetech.org/about/resources/expanding-employment-success-for-people-with-disabilities-2/ (accessed on 17 February 2022). [116]

Fuller, J. et al. (2021), *Hidden Workers: Untapped Talent*, Accenture and Harvard Business School, https://www.hbs.edu/managing-the-future-of-work/Documents/research/hiddenworkers09032021.pdf (accessed on 21 February 2022). [153]

Garcia, D. and S. Sikström (2014), "The dark side of Facebook: Semantic representations of status updates predict the Dark Triad of personality", *Personality and Individual Differences*, Vol. 67, pp. 92-96, https://doi.org/10.1016/J.PAID.2013.10.001. [124]

Garr, S. and C. Jackson (2019), *Diversity & Inclusion Technology: The Rise of a Transformative Market*, RedThread Research and Mercer, https://info.mercer.com/rs/521-DEV-513/images/Mercer_DI_Report_Digital.pdf (accessed on 26 January 2022). [53]

Gaucher, D., J. Friesen and A. Kay (2011), "Evidence that gendered wording in job advertisements exists and sustains gender inequality", *Journal of personality and social psychology*, Vol. 101/1, pp. 109-128, https://doi.org/10.1037/A0022530. [18]

Georgieff, A. and R. Hyee (2021), "Artificial intelligence and employment : New cross-country evidence", *OECD Social, Employment and Migration Working Papers*, No. 265, OECD Publishing, Paris, https://dx.doi.org/10.1787/c2c1d276-en. [85]

Georgieff, A. and A. Milanez (2021), "What happened to jobs at high risk of automation?", *OECD Social, Employment and Migration Working Papers*, No. 255, OECD, Paris, https://www.oecd-ilibrary.org/social-issues-migration-health/what-happened-to-jobs-at-high-risk-of-automation_10bc97f4-en (accessed on 3 March 2022). [88]

Gershgorn, D. (2018), *Companies are on the hook if their hiring algorithms are biased*, Quartz, https://qz.com/1427621/companies-are-on-the-hook-if-their-hiring-algorithms-are-biased/ (accessed on 22 February 2022). [161]

Giermindl, L. et al. (2021), "The dark sides of people analytics: reviewing the perils for organisations and employees", *https://doi.org/10.1080/0960085X.2021.1927213*, https://doi.org/10.1080/0960085X.2021.1927213. [75]

Glassclock, A. (2020), *Chat with Us: How States are Using Chatbots to Respond to the Demands of COVID-19*, National Association of Chief Information Officers, https://www.nascio.org/wp-content/uploads/2020/06/NASCIO_ChatbotsRespondtoCOVID-19.pdf (accessed on 22 December 2022). [26]

Glassdoor (2020), *How to Calculate Cost-Per-Hire*, https://www.glassdoor.co.uk/employers/blog/calculate-cost-per-hire/ (accessed on 19 January 2022). [30]

Goddard, K., A. Roudsari and J. Wyatt (2012), "Automation bias: a systematic review of frequency, effect mediators, and mitigators", *Journal of the American Medical Informatics Association : JAMIA*, Vol. 19/1, pp. 121-127, https://doi.org/10.1136/AMIAJNL-2011-000089. [197]

Gonzalez, M. et al. (2019), ""Where's the I-O?" Artificial Intelligence and Machine Learning in Talent Management Systems", *Personnel Assessment and Decisions*, Vol. 5/3, https://doi.org/10.25035/pad.2019.03.005. [105]

Grassegger, H. and M. Krogerus (2018), *The Data That Turned the World Upside Down*, Vice, https://www.vice.com/en/article/mg9vvn/how-our-likes-helped-trump-win (accessed on 15 April 2022). [178]

Guo, A. et al. (2019), *Toward Fairness in AI for People with Disabilities: A Research Roadmap*, Association for Computing Machinery (ACM), https://doi.org/10.1145/3386296.3386298. [117]

Hardy-Vallée, B. (2019), *Revealing Career Development Opportunities With AI*, SightsIn Plus, https://sightsinplus.com/insight/telant/revealing-career-development-opportunities-with-ai/ (accessed on 22 December 2022). [13]

Harwell, D. (2019), *HireVue's AI face-scanning algorithm increasingly decides whether you deserve the job*, The Washington Post, https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/ (accessed on 15 April 2022). [119]

Hekkala, S. and R. Hekkala (2021), *Integration of artificial intelligence into recruiting young undergraduates: The perceptions of 20-23-year-old students*, IEEE Computer Society, https://doi.org/10.24251/HICSS.2021.021. [34]

Horton, J. (2017), "The effects of algorithmic labor market recommendations: Evidence from a field experiment", *Journal of Labor Economics*, Vol. 35/2, pp. 345-385, https://doi.org/10.1086/689213/SUPPL_FILE/14036DATA.ZIP. [109]

Howard, A., C. Zhang and E. Horvitz (2017), *Addressing bias in machine learning algorithms: A pilot study on emotion recognition for intelligent systems*, IEEE Computer Society, https://doi.org/10.1109/ARSO.2017.8025197. [115]

HR Research Institute (2019), "The 2019 State of Artificial Intelligence in Talent Acquisition", https://www.oracle.com/a/ocom/docs/artificial-intelligence-in-talent-acquisition.pdf (accessed on 18 January 2022). [63]

Hutchinson, A. (2022), *New Survey Shows that Social Media Users are Increasingly Concerned About Data Privacy [Infographic] | Social Media Today*, Social Media Today, https://www.socialmediatoday.com/news/new-survey-shows-that-social-media-users-are-increasingly-concerned-about-d/617548/ (accessed on 9 May 2022). [176]

IAPP (n.d.), *US State Privacy Legislation Tracker*, International Association of Privacy Professionals, https://iapp.org/resources/article/us-state-privacy-legislation-tracker/ (accessed on 21 March 2022). [207]

ILCS (2019), *820 ILCS 42/ Artificial Intelligence Video Interview Act*, https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=4015&ChapterID=68 (accessed on 14 December 2021). [186]

IndustryARC (n.d.), *Artificial Intelligence in Recruitment Market Research Report: Market size, Industry outlook, Market Forecast, Demand Analysis, Market Share, Market Report 2021-2026*, https://www.industryarc.com/Report/19231/artificial-intelligence-in-recruitment-market.html (accessed on 18 January 2022). [70]

Jackson, A. (2016), *Examining factors influencing use of a decision aid in personnel selection*, Kansas State University, https://krex.k-state.edu/dspace/handle/2097/32538 (accessed on 6 April 2022). [192]

Jia, J., G. Jin and L. Wagman (2019), *The short-run effects of GDPR on technology venture investment*, VOXEU CEPR Policy Portal, https://voxeu.org/article/short-run-effects-gdpr-technology-venture-investment (accessed on 3 May 2022). [184]

Joachims, T. et al. (2005), *Accurately interpreting clickthrough data as implicit feedback*, https://doi.org/10.1145/1076034.1076063. [193]

Johansson, J. and S. Herranen (2019), *The application of Artificial Intelligence (AI) in Human Resource Management: Current state of AI and its impact on the traditional recruitment process*, Jönköping University, http://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-44323 (accessed on 21 January 2022). [94]

Jost, J. et al. (2009), "The existence of implicit bias is beyond reasonable doubt: A refutation of ideological and methodological objections and executive summary of ten studies that no manager should ignore", *Research in Organizational Behavior*, Vol. 29, pp. 39-69, https://doi.org/10.1016/J.RIOB.2009.10.001. [43]

Kahn, J. (2021), "HireVue stops using facial expressions to assess job candidates amid audit of its A.I. algorithms", *Fortune*, https://fortune.com/2021/01/19/hirevue-drops-facial-monitoring-amid-a-i-algorithm-audit/ (accessed on 24 November 2021). [211]

Kaibel, C. et al. (2019), "Applicant Perceptions of Hiring Algorithms - Uniqueness and Discrimination Experiences as Moderators", *Academy of Management Proceedings*, Vol. 2019/1, https://doi.org/10.5465/AMBPP.2019.210. [173]

Kang, J. and K. Lane (2010), "Seeing Through Colorblindness: Implicit Bias and the Law by Jerry Kang, Kristin Lane :: SSRN", *UCLA Law Review*, Vol. 58, pp. 465-520, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1633071 (accessed on 14 April 2022). [42]

Keller, S. (2017), *Attracting and retaining the right talent*, McKinsey & Company, https://www.mckinsey.com/business-functions/people-and-organizational-performance/our-insights/attracting-and-retaining-the-right-talent (accessed on 15 April 2022). [136]

Kern, M. et al. (2019), "Social media-predicted personality traits and values can help match people to their ideal jobs", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 116/52, pp. 26459-26464, https://doi.org/10.1073/PNAS.1917942116/SUPPL_FILE/PNAS.1917942116.SAPP.PDF. [127]

Kerr, D. (2020), *Uber's 'racially biased' ratings system hurts nonwhite drivers, lawsuit says*, CNET, https://www.cnet.com/tech/mobile/ubers-racially-biased-ratings-system-hurts-non-white-drivers-lawsuit-says/ (accessed on 17 August 2022). [11]

Kim, J. and W. Heo (2021), "Artificial intelligence video interviewing for employment: perspectives from applicants, companies, developer and academicians", *Information Technology and People*, https://doi.org/10.1108/ITP-04-2019-0173/FULL/PDF. [35]

Kim, P. (2017), "Auditing Algorithms for Discrimination", *University of Pennsylvania Law Review*, Vol. 166, pp. 189-203, https://scholarship.law.upenn.edu/penn_law_review_online/vol166/iss1/10 (accessed on 1 April 2022). [212]

Kim, P. (2017), "Data-Driven Discrimination at Work", *William & Mary Law Review*, Vol. 58/3, pp. 857-936, https://scholarship.law.wm.edu/wmlr/vol58/iss3/4 (accessed on 26 January 2022). [46]

Kim, T. and A. Duhachek (2020), "Artificial Intelligence and Persuasion: A Construal-Level Account", *Psychological Science*, Vol. 31/4, pp. 363-380, https://doi.org/10.1177/0956797620904985. [202]

Kinder, E., M. Jarrahi and W. Sutherland (2019), *Gig Platforms, Tensions, Alliances and Ecosystems*, ACM PUB27 New York, NY, USA, https://doi.org/10.1145/3359314. [141]

Kleinberg, J. and M. Raghavan (2021), “Algorithmic monoculture and social welfare”, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 118/22, https://doi.org/10.1073/PNAS.2018340118/SUPPL_FILE/PNAS.2018340118.SAPP.PDF. [146]

Köchling, A. and M. Wehner (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*, Vol. 13/3, pp. 795-848, https://doi.org/10.1007/S40685-020-00134-W/FIGURES/2. [167]

Korn Ferry (2018), *Korn Ferry Global Survey: Artificial Intelligence (AI) Reshaping the Role of the Recruiter*, https://www.kornferry.com/about-us//press/korn-ferry-global-survey-artificial-intelligence-reshaping-the-role-of-the-recruiter (accessed on 18 January 2022). [62]

Kosinski, M., D. Stillwell and T. Graepel (2013), “Private traits and attributes are predictable from digital records of human behavior”, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 110/15, pp. 5802-5805, https://doi.org/10.1073/PNAS.1218772110/-/DCSUPPLEMENTAL. [177]

Kosinski, M., D. Stillwell and T. Graepel (2013), “Private traits and attributes are predictable from digital records of human behavior”, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 110/15, pp. 5802-5805, https://doi.org/10.1073/PNAS.1218772110/-/DCSUPPLEMENTAL. [122]

Kraiger, K. and J. Ford (1985), “A Meta-Analysis of Ratee Race Effects in Performance Ratings”, *Article in Journal of Applied Psychology*, Vol. 70/1, pp. 56-65, https://doi.org/10.1037/0021-9010.70.1.56. [139]

Krämer, C. and S. Cazes (2022), “Shaping the transition : Artificial intelligence and social dialogue | OECD Social, Employment and Migration Working Papers | OECD iLibrary”, *OECD Social, Employment and Migration Working Papers*, No. 279, OECD, Paris, https://doi.org/10.1787/1bb305a6-en (accessed on 10 November 2022). [2]

Krieger, L. (1995), “The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity”, *Stanford Law Review*, Vol. 47/6, p. 1161, https://doi.org/10.2307/1229191. [41]

Kroll, J. et al. (2017), “Accountable Algorithms”, *University of Pennsylvania Law Review*, Vol. 165, pp. 633-705, https://scholarship.law.upenn.edu/penn_law_review/vol165/iss3/3/ (accessed on 1 April 2022). [213]

Kryscynski, D. et al. (2018), “Analytical abilities and the performance of HR professionals”, *Human Resource Management*, Vol. 57/3, pp. 715-738, https://doi.org/10.1002/HRM.21854. [74]

Lambrecht, A. and C. Tucker (2019), “Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads”, *https://doi.org/10.1287/mnsc.2018.3093*, Vol. 65/7, pp. 2966-2981, https://doi.org/10.1287/MNSC.2018.3093. [152]

Lane, M. and A. Saint-Martin (2021), "The impact of Artificial Intelligence on the labour market: What do we know so far?", *OECD Social, Employment and Migration Working Papers*, No. 256, OECD Publishing, Paris, https://doi.org/10.1787/7c895724-en. [89]

Langer, M., C. König and M. Papathanasiou (2019), "Highly automated job interviews: Acceptance under the influence of stakes", *International Journal of Selection and Assessment*, Vol. 27/3, pp. 217-234, https://doi.org/10.1111/IJSA.12246. [103]

Langer, M. et al. (2020), "Highly automated interviews: applicant reactions and the organizational context", *Journal of Managerial Psychology*, Vol. 35/4, pp. 301-314, https://doi.org/10.1108/JMP-09-2018-0402/FULL/PDF. [106]

Langer, M. and R. Landers (2021), "The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers", *Computers in Human Behavior*, Vol. 123, pp. 1-20, https://doi.org/10.1016/J.CHB.2021.106878. [100]

Laurim, V. et al. (2021), "Computer, whom should I hire? - Acceptance criteria for artificial intelligence in the recruitment process", *Proceedings of the Annual Hawaii International Conference on System Sciences*, Vol. 2020-January, pp. 5495-5504, https://doi.org/10.24251/HICSS.2021.668. [92]

Lawler, E., A. Levenson and J. Boudreau (2004), *HR Metrics and Analytics: Uses and Impacts*, Center for Effective Organizations , Marshall School of Business, University of Southern Califorina, http://www.marshall.usc.edu/ceo (accessed on 8 March 2022). [78]

Lee, M. (2018), "Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management:", *https://doi.org/10.1177/2053951718756684*, Vol. 5/1, https://doi.org/10.1177/2053951718756684. [148]

Lee, M. et al. (2015), *Working with machines: The impact of algorithmic and data-driven management on human workers*, Association for Computing Machinery, https://doi.org/10.1145/2702123.2702548. [142]

Leong, C. (2018), "Technology & recruiting 101: how it works and where it's going", *Strategic HR Review*, Vol. 17/1, pp. 50-52, https://www.emerald.com/insight/content/doi/10.1108/SHR-12-2017-0083/full/pdf (accessed on 19 January 2022). [32]

Lerman, J. (2013), "Big Data and Its Exclusions", *Stanford Law Review Symposium - 2013 - Privacy and Big Data*, Vol. 66, https://www.stanfordlawreview.org/online/privacy-and-big-data-big-data-and-its-exclusions/ (accessed on 21 February 2022). [155]

Levenson, A. (2005), "Harnessing the power of HR analytics", *Strategic HR Review*, Vol. 4/3, pp. 28-31, https://doi.org/10.1108/14754390580000607. [73]

Lewis, N. (2019), *IBM Transforms Its Approach to Human Resources with AI*, SHRM, https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/ibm-transforms-human-resources-ai.aspx (accessed on 27 June 2022). [15]

Li, D., L. Raymond and P. Bergman (2021), "Hiring as Exploration", https://danielle-li.github.io/assets/docs/HiringAsExploration.pdf (accessed on 18 October 2022). [57]

Li, L. et al. (2021), *Algorithmic Hiring in Practice: Recruiter and HR Professional's Perspectives on AI Use in Hiring*, Association for Computing Machinery, Inc, https://doi.org/10.1145/3461702.3462531. [59]

LinkedIn (2016), *Global Recruiting Trends 2017*, https://business.linkedin.com/content/dam/me/business/en-us/talent-solutions/resources/pdfs/linkedin-global-recruiting-trends-report.pdf (accessed on 19 January 2022). [29]

Lisa, A. and V. Talla Simo (2021), *An in-depth study on the stages of AI in recruitment process of HRM and attitudes of recruiters and recruitees towards AI in Sweden*, Umeå University, https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1566298&dswid=4883 (accessed on 3 March 2022). [90]

Lismont, J. et al. (2017), "Defining analytics maturity indicators: A survey approach", *International Journal of Information Management*, Vol. 37/3, pp. 114-124, https://doi.org/10.1016/J.IJINFOMGT.2016.12.003. [80]

Lyell, D. and E. Coiera (2017), "Automation bias and verification complexity: a systematic review", *Journal of the American Medical Informatics Association*, Vol. 24/2, pp. 423-431, https://doi.org/10.1093/JAMIA/OCW105. [199]

Manyika, J. et al. (2015), *A Labor Market That Works: Connecting Talent With Opportunity in the Digital Age*, McKinsey Global Institute, http://www.mckinsey.com/mgi. (accessed on 19 January 2022). [38]

Mcgowan, M. and D. Andrews (2015), "Skill Mismatch and Public Policy in OECD Countries", *OECD Economics Department Working Paper Series*, No. 1210, OECD, Paris, http://www.oecd.org/eco/workingpapers (accessed on 5 February 2019). [6]

McKinsey (2020), *Global survey: The state of AI in 2020*, https://www.mckinsey.com/capabilities/quantumblack/our-insights/global-survey-the-state-of-ai-in-2020 (accessed on 12 April 2022). [64]

McKinsey (2018), *Adoption of AI advances, but foundational barriers remain*, McKinsey & Company, https://www.mckinsey.com/featured-insights/artificial-intelligence/ai-adoption-advances-but-foundational-barriers-remain (accessed on 3 May 2022). [83]

Meyer, J., L. Feinshreiber and Y. Parmet (2003), *Levels of automation in a simulated failure detection task*, https://doi.org/10.1109/ICSMC.2003.1244194. [203]

Minbaeva, D. (2018), "Building credible human capital analytics for organizational competitive advantage", *Human Resource Management*, Vol. 57/3, pp. 701-713, https://doi.org/10.1002/HRM.21848. [71]

Mirowska, A. and L. Mesnet (2021), "Preferring the devil you know: Potential applicant reactions to artificial intelligence evaluation of interviews", *Human Resource Management Journal*, https://doi.org/10.1111/1748-8583.12393. [91]

Misuraca, G. and C. van Noordt (2020), *Overview of the use and impact of AI in public services in the EU*, Publications Office of the European Union, Luxembourg, https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/ai-watch-artificial-intelligence-public-services (accessed on 11 March 2022). [183]

Mohlmann, M. and L. Zalmanson (2017), “Hands on the Wheel: Navigating Algorithmic Management and Uber Drivers’ Autonomy”, *ICIS 2017 Proceedings*, https://aisel.aisnet.org/icis2017/DigitalPlatforms/Presentations/3 (accessed on 15 February 2022). [143]

Moore, P. (2020), *Data subjects, digital surveillance, AI and the future of work*, European Union, Brussels, https://www.europarl.europa.eu/RegData/etudes/STUD/2020/656305/EPRS_STU(2020)656305_EN.pdf (accessed on 28 April 2022). [97]

Mori, K. and M. Haruno (2021), “Differential ability of network and natural language information on social media to predict interpersonal and mental health traits”, *Journal of Personality*, Vol. 89/2, pp. 228-243, https://doi.org/10.1111/JOPY.12578. [128]

Morrison & Foerster (n.d.), *The Risks of AI-Driven HR Tools for Data Protection and Privacy*, https://media2.mofo.com/documents/211001-risks-ai-driven-hr-tool.pdf (accessed on 9 May 2022). [219]

Morrison, R. (2022), *AI is better at recruitment than humans, but still lacks the personal touch*, Tech Monitor, https://techmonitor.ai/technology/ai-and-automation/ai-recruitment-bias-humans (accessed on 21 October 2022). [149]

Mortensen, D. and C. Pissarides (1999), “New developments in models of search in the labor market”, in *Handbook of Labor Economics*, Elsevier, https://doi.org/10.1016/S1573-4463(99)30025-0. [4]

Mosier, K. et al. (2009), “Aircrews and Automation Bias: The Advantages of Teamwork?”, *The International Journal of Aviation Psychology*, Vol. 11/1, pp. 1-14, https://doi.org/10.1207/S15327108IJAP1101_1. [205]

Mosier, K. et al. (1998), “Automation Bias: Decision Making and Performance in High-Tech Cockpits”, *The International Journal of Aviation Psychology*, Vol. 8/1, pp. 47-63, https://lskitka.people.uic.edu/AutomationBias1998.pdf (accessed on 7 April 2022). [194]

Mueller, B. (2021), *AI Act Would Cost the EU Economy €31 Billion Over 5 Years, and Reduce AI Investments by Almost 20 Percent, New Report Finds*, Center for Data Innovation, https://datainnovation.org/2021/07/ai-act-would-cost-the-eu-economy-e31-billion-over-5-years-and-reduce-ai-investments-by-almost-20-percent-new-report-finds/ (accessed on 3 May 2022). [185]

NCSL (n.d.), *State Social Media Privacy Laws*, National Conference of State Legislatures, https://www.ncsl.org/research/telecommunications-and-information-technology/state-laws-prohibiting-access-to-social-media-usernames-and-passwords.aspx (accessed on 21 March 2022). [208]

Newman, D., N. Fast and D. Harmon (2020), “When eliminating bias isn’t fair: Algorithmic reductionism and procedural justice in human resource decisions”, *Organizational Behavior and Human Decision Processes*, Vol. 160, pp. 149-167, https://doi.org/10.1016/J.OBHDP.2020.03.008. [170]

Nguyen, L. and D. Gatica-Perez (2016), “Hirability in the Wild: Analysis of Online Conversational Video Resumes”, *IEEE Transactions on Multimedia*, Vol. 18/7, pp. 1422-1437, https://doi.org/10.1109/TMM.2016.2557058. [111]

Nocker, M. and V. Sena (2019), “Big Data and Human Resources Management: The Rise of Talent Analytics”, *Social Sciences*, Vol. 8/10, https://doi.org/10.3390/SOCSCI8100273. [82]

OECD (2022), *OECD Framework for the Classification of AI systems*, OECD, Paris, https://doi.org/10.1787/20716826 (accessed on 1 August 2022). [7]

OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris, https://dx.doi.org/10.1787/eedfee77-en. [214]

OECD (2018), *“The Work”, AI Job Recommendation Service Using the National Job Information Platform*, Observatory of Public Sector Innovation Observatory of Public Sector Innovation, https://oecd-opsi.org/innovations/the-work/ (accessed on 19 January 2022). [36]

Oswald, F. et al. (2020), “Big Data in Industrial-Organizational Psychology and Human Resource Management: Forward Progress for Organizational Research and Practice”, *Annual Review of Organizational Psychology and Organizational Behavior*, Vol. 7, pp. 505-533, https://doi.org/10.1146/ANNUREV-ORGPSYCH-032117-104553. [179]

Ötting, S. and G. Maier (2018), “The importance of procedural justice in Human–Machine Interactions: Intelligent systems as new decision agents in organizations”, *Computers in Human Behavior*, Vol. 89, pp. 27-39, https://doi.org/10.1016/J.CHB.2018.07.022. [175]

Parasuraman, R. and D. Manzey (2010), “Complacency and bias in human use of automation: an attentional integration”, *Human factors*, Vol. 52/3, pp. 381-410, https://doi.org/10.1177/0018720810376055. [198]

Parasuraman, R. and V. Riley (1997), “Humans and Automation: Use, Misuse, Disuse, Abuse”, *https://doi.org/10.1518/001872097778543886*, Vol. 39/2, pp. 230-253, https://doi.org/10.1518/001872097778543886. [196]

Polli, F. (2019), *Using AI to Eliminate Bias from Hiring*, Harvard Business Review, https://hbr.org/2019/10/using-ai-to-eliminate-bias-from-hiring (accessed on 6 May 2022). [218]

PWC (2017), *Artificial Intelligence in HR: A No-brainer*, https://www.pwc.nl/nl/assets/documents/artificial-intelligence-in-hr-a-no-brainer.pdf (accessed on 18 January 2022). [61]

Raghavan, M. et al. (2019), *Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices*, Elsevier BV, https://doi.org/10.2139/SSRN.3408010. [54]

Raji, I. and J. Buolamwini (2019), *Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products*, Association for Computing Machinery, Inc, https://doi.org/10.1145/3306618.3314244. [114]

Raub, M. (2018), “Bots, Bias and Big Data: Artificial Intelligence, Algorithmic Bias and Disparate Impact Liability in Hiring Practices”, *Arkansas Law Review*, Vol. 71/2, https://scholarworks.uark.edu/alr/vol71/iss2/7 (accessed on 14 January 2022). [157]

Raymond, J. (2013), “Most of us are biased”, *Nature*, Vol. 495, pp. 33-34, https://doi.org/10.1038/495033a. [44]

Rhue, L. (2018), “Racial Influence on Automated Perceptions of Emotions”, *SSRN Electronic Journal*, University of Maryland, https://doi.org/10.2139/SSRN.3281765. [113]

Rosenbaum, E. (2019), *IBM AI can predict with 95 percent accuracy which employees will quit*, CNBC, https://www.cnbc.com/2019/04/03/ibm-ai-can-predict-with-95-percent-accuracy-which-employees-will-quit.html (accessed on 27 June 2022). [14]

Salvi del Pero, A., P. Wyckoffi and A. Vourc'h (2022), "Using Artificial Intelligence in the workplace : What are the main ethical risks?", *OECD Social, Employment and Migration Working Papers*, No. 273, https://doi.org/10.1787/840a2d9f-en (accessed on 13 November 2022). [1]

Sánchez-Monedero, J., L. Dencik and L. Edwards (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", *FAT* 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 458-468, https://doi.org/10.1145/3351095.3372849. [23]

Sarpong, S. and D. Rees (2014), "Assessing the effects of 'big brother' in a workplace: The case of WAST", *European Management Journal*, Vol. 32/2, pp. 216-222, https://doi.org/10.1016/J.EMJ.2013.06.008. [98]

Schwartz, H. et al. (2013), "Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach", *PLOS ONE*, Vol. 8/9, https://doi.org/10.1371/JOURNAL.PONE.0073791. [123]

Schwartz, R. et al. (n.d.), "Towards a Standard for Identifying and Managing Bias in Artificial Intelligence", *NIST Special Publication*, Vol. 1270, p. 1270, https://doi.org/10.6028/NIST.SP.1270. [168]

Schwellnus, C. et al. (2019), "Gig economy platforms: Boon or Bane?", *Economics Department Working Papers*, No. 1550, OECD, Paris, https://doi.org/10.1787/fdb0570b-en (accessed on 19 January 2022). [37]

SHRM (2016), *2016 Human Capital Benchmarking Report*, Society for Human Resource Management, https://www.shrm.org/hr-today/trends-and-forecasting/research-and-surveys/pages/2016-human-capital-report.aspx (accessed on 19 January 2022). [31]

Simonite, T. (2018), *AI Is the Future—But Where Are the Women?*, https://www.wired.com/story/artificial-intelligence-researchers-gender-imbalance/ (accessed on 22 February 2022). [160]

Skitka, L., K. Mosier and M. Burdick (2000), "Accountability and automation bias", *International Journal of Human Computer Studies*, Vol. 52/4, pp. 701-717, https://doi.org/10.1006/IJHC.1999.0349. [200]

Skitka, L., K. Mosier and M. Burdick (1999), "Does automation bias decision-making?", *International Journal of Human-Computer Studies*, Vol. 51/5, pp. 991-1006, https://doi.org/10.1006/IJHC.1999.0252. [195]

Skitka, L. et al. (2000), "Automation bias and errors: are crews better than individuals?", *The International Journal of Aviation Psychology*, Vol. 10/1, pp. 85-97, https://doi.org/10.1207/S15327108IJAP1001_5. [206]

Smith, A. and M. Anderson (2017), *Americans and Automation in Everyday Life*, Pew Research Center, https://www.pewresearch.org/internet/2017/10/04/automation-in-everyday-life/ (accessed on 11 February 2022). [99]

Stauffer, J. and M. Buckley (2005), “The existence and nature of racial bias in supervisory ratings”, *The Journal of applied psychology*, Vol. 90/3, pp. 586-591, https://doi.org/10.1037/0021-9010.90.3.586. [138]

Stolz, R. (2018), *Using ‘cognitive HR’ systems to optimize talent deployment*, Employee Benefit News, https://www.benefitnews.com/news/using-cognitive-hr-systems-to-optimize-talent-deployment (accessed on 27 June 2022). [12]

Suen, H., M. Chen and S. Lu (2019), “Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes?”, *Computers in Human Behavior*, Vol. 98, pp. 93-101, https://doi.org/10.1016/J.CHB.2019.04.012. [174]

Sühr, T., S. Hilgard and H. Lakkaraju (2020), *Does Fair Ranking Improve Minority Outcomes? Understanding the Interplay of Human and Algorithmic Biases in Online Hiring*, Association for Computing Machinery, Inc, https://doi.org/10.48550/arxiv.2012.00423. [56]

Tambe, P., P. Cappelli and V. Yakubovich (2019), “Artificial Intelligence in Human Resources Management: Challenges and a Path Forward:”, *California Management Review*, Vol. 61/4, pp. 15-42, https://doi.org/10.1177/0008125619867910. [67]

Tatman, R. (2017), *Gender and Dialect Bias in YouTube’s Automatic Captions*, Association for Computational Linguistics (ACL), https://doi.org/10.18653/v1/W17-1606. [118]

Textmetrics (n.d.), *Discover how Randstad increased their conversion by double digits, using Textmetrics technology*, https://www.textmetrics.com/customer-story-with-randstad/ (accessed on 30 August 2022). [17]

Thomé, J. (n.d.), *How to Optimize Your Resumé for AI Hiring Systems*, https://www.ellevatenetwork.com/articles/10725-how-to-optimize-your-resume-for-ai-hiring-systems (accessed on 17 February 2022). [144]

Tobin, A. and J. Merrill (2018), *Facebook Is Letting Job Advertisers Target Only Men*, ProPublica, https://www.propublica.org/article/facebook-is-letting-job-advertisers-target-only-men (accessed on 29 March 2022). [216]

Uggerslev, K., N. Fassina and D. Kraichy (2012), “Recruiting Through the Stages: A Meta-Analytic Test of Predictors of Applicant Attraction at Different Stages of the Recruiting Process”, *Personnel Psychology*, Vol. 65/3, pp. 597-660, https://doi.org/10.1111/J.1744-6570.2012.01254.X. [104]

Ulfert, A., C. Antoni and T. Ellwart (2022), “The role of agent autonomy in using decision support systems at work”, *Computers in Human Behavior*, Vol. 126, pp. 1-13, https://doi.org/10.1016/J.CHB.2021.106987. [87]

Unsworth, N. et al. (2015), “Is Playing Video Games Related to Cognitive Abilities?”, *Psychological Science*, Vol. 26/6, pp. 759-774, https://doi.org/10.1177/0956797615570367. [134]

Upadhyay, A. and K. Khandelwal (2018), “Applying artificial intelligence: implications for recruitment”, *Strategic HR Review*, Vol. 17/5, pp. 255-258, https://doi.org/10.1108/SHR-07-2018-0051. [95]

van den Broek, E., A. Sergeeva and M. Husyman (2019), *Hiring algorithms: An ethnography of fairness in practice*, Association for Information Systems, https://research.vu.nl/en/publications/hiring-algorithms-an-ethnography-of-fairness-in-practice (accessed on 6 May 2022). [49]

van Esch, P., J. Black and J. Ferolie (2019), "Marketing AI recruitment: The next phase in job application and selection", *Computers in Human Behavior*, Vol. 90, pp. 215-222, https://doi.org/10.1016/J.CHB.2018.09.009. [165]

van Es, K., D. Everts and I. Muis (2021), "Gendered language and employment Web sites: How search algorithms can cause allocative harm", *First Monday*, Vol. 26/8, https://doi.org/10.5210/FM.V26I8.11717. [158]

Van Iddekinge, C. et al. (2013), "Social Media for Selection? Validity and Adverse Impact Potential of a Facebook-Based Assessment:", *Journal of Management*, Vol. 42/7, pp. 1811-1835, https://doi.org/10.1177/0149206313515524. [130]

Vanmeirhaeghe, H. (2021), *AI Recruitment: De Toekomst van Human Resources?*, University of Ghent, Ghent, https://libstore.ugent.be/fulltxt/RUG01/002/839/996/RUG01-002839996_2020_0001_AC.pdf (accessed on 21 January 2022). [93]

Verhagen, A. (2021), "Opportunities and drawbacks of using artificial intelligence for training", *OECD Social, Employment and Migration Working Papers*, No. 266, OECD Publishing, Paris, https://doi.org/10.1787/22729bd6-en. [25]

von Lewinski, K. and R. de Barros Fritz (2022), "Chapter 18: Legitimacy of artificial intelligence in human resources - the legal framework for using artificial intelligence in human resource management", in Strohmeier, S. (ed.), *Handbook of Research on Artificial Intelligence in Human Resource Management*, Elgar, https://www.elgaronline.com/view/edcoll/9781839107528/9781839107528.00030.xml (accessed on 10 June 2022). [189]

Wang, R., F. Harper and H. Zhu (2020), "Factors Influencing Perceived Fairness in Algorithmic Decision-Making: Algorithm Outcomes, Development Procedures, and Individual Differences", *Conference on Human Factors in Computing Systems - Proceedings*, pp. 1-14, https://doi.org/10.1145/3313831.3376813. [172]

WEF (2020), *How can AI recruitment systems reduce hiring bias?*, The Jobs Reset Summit, World Economic Forum, https://www.weforum.org/agenda/2020/10/ai-hiring-tools-bias-recruiting-hiring-diversity-fairness-equality/ (accessed on 21 January 2022). [39]

Wehner, M., A. Köchling and J. Warkocz (2021), *Can I perform? Affective responses to algorithmic decision-making in the recruitment process*, Academy of Management Briarcliff Manor, NY 10510, https://doi.org/10.5465/AMBPP.2021.15205ABSTRACT. [102]

Will, P., D. Krpan and G. Lordan (2022), "People versus machines: introducing the HIRE framework", *Artificial Intelligence Review*, pp. 1-30, https://doi.org/10.1007/S10462-022-10193-6/TABLES/9. [107]

Yam, J. and J. Skorburg (2021), "From human resources to human rights: Impact assessments for hiring algorithms", *Ethics and Information Technology 2021 23:4*, Vol. 23/4, pp. 611-623, https://doi.org/10.1007/S10676-021-09599-7. [52]

Zerilli, J. et al. (2018), "Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard?", *Philosophy & Technology*, Vol. 32/4, pp. 661-683, https://doi.org/10.1007/S13347-018-0330-6. [181]

Zhang, H., S. Feinzig and H. Hemmingham (2018), *Making Moves: Internal Career Mobility and the Role of AI*, IBM Talent Management Solutions, https://www.ibm.com/downloads/cas/X2O7V1OD (accessed on 22 December 2022). [16]

Zwitter, A. (2014), "Big Data ethics:", *Big Data & Society*, pp. 1-6, https://doi.org/10.1177/2053951714559253. [187]

Annex A. List of Interviewees

Organisation Type and Number	Job Title
Disability Ethical AI? Alliance (DEAI), Civil Society	Susan Scott-Parker, Founder
Pymetrics	Sara Kasir, Principal Public Policy and Research
Manufacturing Firm	HR Adviser
PES 1	Data Scientist Lead
PES 2	External Relations, AI and Data
PES 2	Analyst
PrES 1	Group SVP Head of Data & AI
PrES 1	AI Product Manager
PrES 1	Vice President Digital Product Marketing
PrES 1	Senior Public Affairs Manager
PrES 2	Senior Legal Counsel
PrES 2	Head of Digital Strategy
PrES 3	Global Privacy Director
LinkedIn	Joaquin Quinonero Candela, Technical Fellow AI
LinkedIn	Senior Policy Manager
LinkedIn	Legal Director
Job Board Operator	Head of Machine Learning
Job Board Operator	Director of Engineering (Data Organisation)
Consulting Company	Manager Digital Ethics
Consulting Company	Workforce Strategy & Transformation Leader
Consulting Company	HR Innovation Strategist
Consulting Company	A Digital Transformation Leader
Consulting Company	Managing Director for Workforce Transformation

Annex B. Interview Guides

Public Employment Services

The following questions were used as a basis for discussions PES headquarters (e.g. AI strategists, IT department).

Use of AI technologies

Q1. For what kind of services is the PES already using AI? For example:

Regarding jobseekers

- Helping to find information on the support measures they are entitled to
- Helping jobseekers to fill in their profile (e.g. based on the ESCO taxonomy)
- Profiling and early identification of hard-to-employ jobseekers
- Helping jobseekers to find vocational education and training programmes
- Other services – please specify

Regarding employers

- Helping employers find information on support measures they are entitled to
- Helping employers to fill in their job ad (e.g. based on the ESCO taxonomy)
- Early identification of hard-to-fill vacancies
- Early identification of potential new vacancies
- Matching jobseekers with vacancies
- Other services – please specify

Regarding caseworkers in PES agencies

- Helping caseworkers in the execution of repetitive and/or administrative tasks
- Helping caseworkers to develop individual action plans (i.e. support for decision making)
- Helping caseworkers to follow up jobseekers throughout their individual action plan
- Helping caseworkers to monitor job search
- Other services – please specify

Q2. In terms of further development and deployment of AI, what has been planned for the coming years?

Decision-making process

Q3. To start with, where do ideas for AI needs come from, and how are they collected? (E.g. innovation labs, bottom-up vs top-down approaches?)

Q4. Who takes the decision to adopt an AI technology?

Q5. Who is consulted? E.g. PES staff, trade unions, employer associations, any jobseeker representatives?

Q6. What are the main issues discussed during the decision-making process?

Development process

Q7. Are AI technologies developed:

- Mainly in-house? If so, do you also hire external consultants?
- Mainly by contracting out to external organisations. If so:
  - Are these organisations public bodies or private firms?
  - What are the main criteria for selecting a contractor?
  - Are there any off-the-shelf packages?
- It varies from one project to another

Q8. Are PES experts or/and users representatives (e.g. PES staff, jobseekers, employers) involved? If so, at what stage of the development process:

- Early on, for developers to get a good understanding of users' needs and expectations
- Later on, for developers to collect feedback on the approach they are following
- How frequently?

Q9. Are the objectives and design of an AI technology evolving throughout this process of consultation and/or cooperation (e.g. agile development methodologies)?

Evaluation process

Q10. Do you run any pilots to evaluate and fine-tune new AI technologies in development? If so, what did you learn from these pilots?

Q11. Are there any quality standards or norms for certifying a new AI technology before it can be adopted by PES agencies? If so:

- By whom have these standards been developed?
- What are their main objectives?
- Who is in charge of delivering the corresponding certificates?
- What is the certification process?

Q12. Are there regular impact evaluations of the AI technologies used by PES agencies? For example:

- Do you ask users to rate/evaluate the technology? If so, what questions do you ask (e.g. satisfaction, usefulness, relevance), and what do these users' feedbacks show?
- What else do you measure to monitor the impact of the technology? If so, what kind of indicators/measures do you use and what do they shwo?
- How are these results (user rating and other measurements) then used and taken into account? Is the technology regularly fine-tuned?
- Do you conduct any cost-benefit analysis? If so, how do you measure costs? How do you measure benefits? What do these cost-benefit analyses show?

Ethics and governance of AI technologies

Q13. To start with, is there any formal or agreed definition of:

- "AI Ethics" in the PES context? If so, please specify
- "AI Governance" in the PES context? If so, please specify

Q14. Is there a specific board or committee within the PES in charge of AI ethics and governance? If so, what are its main objectives and activities?

Q15. Are there any guidelines, regulations, or legal provisions to ensure that AI technologies are used in an ethical way? For example:

- To protect users' data privacy
- To ensure a right to information when an AI technology is used
- To ensure that any decision driven by an AI technology is overseen by a PES employee
- To ensure a right to explanation of any decision driven by an AI technology
- To ensure a right to challenge any decision driven by an AI technology
- Other important aspects – please specify.

Q16. Are you receiving questions, requests or complains from users on the various aspects we just mentioned? If so, on what particular aspects, and how frequently?

Information and training for PES caseworkers

Q17. Are there general guidelines, information materials or training programmes for PES staff:

- To help them understand what AI technologies can do and cannot do
- To explain the advantages and drawbacks of AI technologies
- To provide concrete recommendations on how they should use AI technologies
- Other – please specify.

Closing questions

Is there anything else you would like to share about the development, implementation and use of AI technologies for delivering PES services?

In the second part of our case study, we would like to understand how AI technologies work in practice, including the implications for PES staff in agencies and PES clients. To this end, we would like to interview project managers and employees who are familiar with those technologies and use them regularly.

Amongst the services already using AI that you mentioned at the beginning of the interview [question A1]:

- On which ones, have PES agencies gained most experience and expertise?
- For each of the services you just cited, could you please indicate the name of a project manager or team leader we could contact and interview?

Private Employment Services

The following questions were used as a basis for discussions Private Employment Services

Basic information about you

Q01. Please let us know about you and your role at the company
- Job title
- Main tasks/responsibilities

Your company's use of AI

Q02. Does your company use AI in its delivery of services?
- If so, please identify AI applications your company has implemented

Applications implemented	**Yes/No**
Optimise job descriptions	
Identification of vacancies: • Assisting job seekers with finding vacancies • Identifying vacancies in general	
Parsing candidate CVs	
Screening	
Background checks	
Skill assessments	
Candidate evaluation, including for behavioural/cultural fit	
Programmatic advertising	
Aide interview scheduling	
Interviewing	
Automated analysis of biometric or video data	
Matching of candidates to vacancies	
Improve communication with candidates/firms	
Identifying skills gaps	
Career pathing: advise job seekers on job paths and training	
Systems to rate current employees	
Other (please specify)	

Basic information about the AI technology

Of the AI technologies that you have mentioned, the next questions related to one specific technology in particular.

Q03. What does [the technology] do? *[If possible, try to elicit a non-technical response.]*
- Specific tasks?
- ➢ In what sense does [the technology] play a core role in business operations?

Q04. What was the business rationale behind the introduction of [the technology]?
- To improve an existing process? → How (e.g., speed, scale, quality)?
- To free up staff time to focus on different tasks? → How? Which tasks?
- To reduce costs → How?

Q05. At what stage of maturity is [the technology] now?
- Used selectively or in multiple activities across the company?
- Plans to change/expand/develop further?
- Type of maintenance [the technology] requires?

- Was [the technology] developed in-house or purchased from a third party?

Q06. How do workers interact with [the technology]?
- Setting of parameters?
- Signing off on the outputs or recommendations of [the technology]?

Q07. How has [the technology] impacted jobs within your company?
- Are there certain tasks that are no longer performed by humans?
- What do humans do with the time freed up?
- Changes to overall headcount?
- Changed job descriptions? Which?
- Created new roles? Which?
- Discontinued some roles? Which?

Benefits and risks of the AI technology

Q08. What are [the technology's] main benefits? *[This question has multiple sub-parts; it will ask about the benefits to different stakeholders.]*
- For job seekers → faster matching/less time spent unemployed/searching for jobs, greater satisfaction with services, more/better job opportunities, better/more timely communication/feedback, simplified application processes?
- ➢ Do these benefits accrue to any group of job seekers, in particular?
- For client firms → faster recruitment, lower costs, better matches, better retention, better representation of different groups?
- For your company → cost savings, efficiencies, new/better services, market share, profits, changes to staff headcount, slowing of new hires, well being?

Q09. Have the benefits of [the technology] been measured?
- Thinking of available metrics, is there evidence of how [the technology] has improved services, e.g., number of people matched to vacancies, number of job seekers, number of vacancies, time to match, satisfaction of employers and job seekers, job retention rates, profits, costs, market share (precise figures or estimates)?
- ➢ In the absence of available metrics, what is your impression of [the technology's] impact?

Q10. How would you describe job seekers' attitudes towards AI? Positive? Negative?
- Are you aware of evidence (e.g., gathered via surveys)?
- ➢ Has your company sought to influence job seekers' attitudes towards AI or experiences of [the technology], such as through tutorials on effective use? How?

Q11. What are [the technology's] main risks or drawbacks? *[As above, this question has multiple sub-parts; it will ask about the risks/drawbacks to different stakeholders.]*
- For job seekers → concerns over whether [the technology] offers a genuine improvement in finding a job, loss of human touch/advice/feedback through the search process, privacy, transparency, technological fluency/ability to navigate the tools?
- ➢ Do these risks impact any group of job seekers, in particular?
- For client firms → concerns over whether [the technology] offers a genuine improvement in incoming staff, loss of human touch?
- For your company → concerns over reputation/bad press?

Q12. Thinking about your personal experience of [the technology], what have been the key changes?
- Which tasks do you no longer perform, and which new tasks do you perform?
  - Automation of boring/repetitive tasks?
  - More interesting work? Less interesting work? Of what kind?
- Are your skills used differently? Are new skills demanded?
- Better decisions (given more/better information)?
- Less decision-making power? Is that good or bad?
- Changes to overall workload? Work intensity?

- How has [the technology] impacted your wellbeing?
    - Stress/pressure, anxiety and mental health?
    - General job satisfaction?
    - Empowerment/control over your work?
- How would you describe your attitude towards [the technology] today? Positive? Negative?
    - Has this changed over time? How?
    - How would you feel about more AI being used in your workplace? Would you have any concerns? In what way?

Barriers to adoption of AI used to improve services

Q13. In your view, what are the most significant barriers to the adoption of AI?
- Within your company:
    - Have staff resisted the introduction of AI technologies? How was this managed?
    - Have client firms resisted the introduction of AI technologies? How was this managed?
- Within the private employment services industry: Reputational concerns? Regulation? Poor digital infrastructure? Lack of data or data quality concerns? Lack of skills to implement/use the technologies?
- ➢ How could/should these issues be overcome?

Data issues

Q14. Have there been discussions about your company's AI use raising any of the following concerns?
- Bias or discrimination regarding certain groups of job seekers?
- Lack of explainability?
- ➢ What measures have been taken?
    - How does your company ensure the quality of data used?
    - Are AI tools audited? With what frequency?

Government policy and regulation

Q15. Have government policy and regulation had any impact on the decision to implement [the technology], how to implement it or any modifications that have been made along the way?
- Does policy and regulation affect the degree of human interaction with [the technology] (e.g., sign-off on decisions recommended)?
- Legislation around personal data protection or algorithmic decision-making?
- Does your company's use of AI vary by country according to regulations in place? How so?

Q16. What are your views on the proposed EU AI regulations?
- Are there modifications that you would like to see?
- Is there any other government policy or regulation you would like to see?

Q17. What kind of government policy and regulation would you like to see? What should policymaking avoid?

Closing Questions

Q18. Has your company experienced any unanticipated challenges or benefits with [the technology]? What were they?
- Were the challenges resolved? → How?
- Anything that you would have done differently? Any lessons learned?
- What were key factors for success?

Q19. Is there anything else you would like to share?

Other

The following questions were used as a basis for discussions with: developers, users, online recruitment platforms and others. The questions were slightly adjusted depending on the interviewee.

1. Please describe the AI technologies that your firm develops/uses. What stages of the recruitment process do they support and what tasks do they automate?
2. What is the nature of the human-AI interaction? Do these tools substitute for humans, or complement them? What skills do humans need to use these technologies?
3. What are the main benefits of using these technologies in the recruitment process? E.g. efficiencies, cost savings, quality of matching, etc. Any data to back this up?
4. What data are used in training the technologies? What outcomes do they try to predict and how were these outcomes chosen? What data do they collect and assess, and what are the outputs of the tools? How explainable are the decisions arrived at? To what extent are the tools customised to each client?
5. How accurate are these tools, compared to humans?
6. How is potential bias addressed in the tools/models? Can the tool help improve diversity in the firm and, if so, how?
7. Where is the technology heading next? What are the main barriers to further develop these technologies? What are the main barriers to adoption?
8. Have government policy and regulation had any impact on the development/adoption of the technology? How are different regulations across different markets dealt with? What would you like to see from government?