



THE UNIVERSITY OF
MELBOURNE

The Policy Lab

Research Paper

Gendered Bots? Bias in the use of Artificial Intelligence in Recruitment

2020

Sheilla Njoto

Contents

Abstract	2
Introduction	3
Artificial Intelligence and the expansion of inequality.....	5
How do AI algorithms work?.....	5
Profit-driven algorithmic judgements: the replica of historical biases	7
Generalisation and microtargeting: discrimination?.....	7
The problem of misrepresentation	9
Statistical bias paving the way for systemic bias	11
Algorithmic biases in determining hiring decisions	13
Jobs ads: maximum engagement through microtargeting	14
Job sorting: translating past history into future outcomes	15
Algorithmic bias against women: the mirror of societal discrimination	19
Correlational bias in hiring: women's future in the hands of men	19
The underrepresentation of women in labour and data	20
Gendered roles, gendered algorithmic judgements	21
Identical but gendered: hiring algorithms in assessing candidates	23
Stage One: James and Jane study	23
Stage Two: online job profiling study	24
Stage Three: motherhood and gender study	25
Stage Four: gendered keywords study	26
Prototyping a world of equality	31
Re-defining discrimination: public policy, regulations, laws and beyond .	28
Making public policy matter	28
Redefining the praxis of transparency and accountability	29
A call for further research	30
Appendices	32
Appendix 1	32
Appendix 2	34
Bibliography	36

Declaration

This is to certify that:

- i. This paper comprises only my original work except where indicated,
- ii. Due acknowledgement has been made in the text to all other material used.

This report was completed as part of the Executive Internship, The University of Melbourne as a contribution to The Future of Women @ Work series.



Sheilla Njoto is a Master of Public Policy and Management student.
sheillanjoto@gmail.com



Abstract

This paper examines how the adoption of hiring algorithms expands gender inequality in the labour force. It hypothesises that, given its nature in predicting the future based on historical data, hiring algorithms exhibit the risk of discriminating women. Its proneness in repeating, if not expanding, societal bias is predominantly geared by its propensity in blindly feeding on (1) data misrepresentation, (2) correlational errors, and (3) the limitation of datasets. The study presents that, firstly, despite the identical qualifications, skills and experience, hiring algorithms rank male candidates higher than female. This includes 'passive' submissions on online candidate profiling. Secondly, despite the overrepresentation of women in parenthood and non-traditional workforce, hiring algorithms discriminate both male and female candidates with parental leave in comparison to those without. Thirdly, it reveals that hiring algorithms are significantly more prone to conceiving gender discrimination in assessing gender-based keyworded resumés compared to the entirely identical resumés. This paper has demonstrated that the rise of digitalisation should redefine the meaning of 'fairness', 'discrimination' and 'accountability'. Despite the seriousness of these problems, however, the lack of cross-disciplinary study in this particular issue pertains. This paper's contentions are a mere reprise of arguments that offer complex theories. It has sought to start a new conversation about the acute problems faced by women.

Introduction

This paper examines how the adoption of hiring algorithms expands gender inequality in the labour force. It hypothesises that, given its nature in predicting the future based on historical data, hiring algorithms exhibit the risk of discriminating women. Its proneness in repeating, if not expanding, societal bias is predominantly geared by its propensity in blindly feeding on (1) data misrepresentation, (2) correlational errors, and (3) the limitation of datasets. This is referred to in this paper as the algorithmic bias.¹ It assumes that, firstly, the future of an *individual* woman is dictated by the historical data of *all* women in aggregate; and secondly, the past and current gender inequality may lead to unequal opportunities for women in the future.

It is to note that the basis for focusing on the market-oriented hiring algorithms lies on the premise that it is largely centred around the principles of *meritocracy*, which means that candidates' contribution (qualifications, skills and experiences) should equal to their rewards (hiring outcomes) (Adams 1965; Homans 1961; Greenberg 1987); rather than intentionally applying 'positive discrimination' or 'affirmative action' for reparative purposes. Hiring algorithms appear as a promising response to the demand of meritocratic judgements. It is no doubt that the upsurge of algorithmic decisions has spurred a considerable rise of efficient decision-making. However, key literature has indicated that algorithms are not entirely exempt from bias.

This paper investigates the implications of this for the representation of women in the labour force. It seeks to answer the questions: (1) given the underrepresentation of women in global and Australia's employment data, would algorithms discriminate against women? (2) given algorithms' proneness to drawing correlational biases between groups of minorities with their history, would this severely disadvantage women? (3) as men stay longer in the workforce than women, usually due to motherhood, would algorithms assume that men are more likely to be successful in job candidacy? (4) how would the different use of language between male and female candidates affect hiring predictions?

In seeking to excavating this issue, this paper seeks to excavate the issue by expounding on theoretical inferences and empirical evidence in six chapters. Chapter One defines Artificial Intelligence and introduces its relation to Big Data. It describes the significance of Big Data and social data to deep learning. By drawing on scholars such as Zuboff and Galloway, Chapter Two will introduce the concept of data commoditisation and its tolerance towards the risk of algorithmic inaccuracies. It will investigate how marketised microtargeting algorithms lead to

¹ In this paper, 'bias' is defined as the decision-making pattern that subjectively benefits a certain group over the others (Dickson *et al.* 2018).

over-generalisation, discrimination and misrepresentation. Chapter Three offers an in-dept analysis on hiring algorithmic biases and its corollary towards discriminatory judgements. In refining the scope, this chapter will begin by identifying the targeted purpose of hiring algorithms in recruitment. Following this, it will analyse how hiring algorithms are prone to conceiving discrimination and bias, especially in microtargeting job advertisements and candidates' success predictions. Alluding to the theoretical presuppositions, Chapter Four and Five concentrate its analysis on the impact of algorithmic biases specifically on women using both empirical and theoretical observations. It presents a body of experiment that is designed to manipulate hiring algorithms using variables that are tailored according to gender, parenthood and gendered keywords. Chapter Six will conclude by indicating legal and social inferences and denoting a call for further research.

This paper is predominantly informed by theoretical analyses. Drawing on the existing research on algorithmic biases and correlational errors, this paper draws logical and theoretical inferences on its impact towards gender inequality. Following this, it will conduct a body of experiments that are rigorously designed to manipulate hiring algorithms using fictitious resumés that are tailored specifically according to gender, parenthood and gendered keywords. Given the intricate nature of AI algorithms, however, this paper identifies several limitations to the study. Firstly, as AI improves its own algorithms at task, it is rather impossible to understand specifically *how* and *why* algorithms make certain judgements and predictions (Faragher 2019). It implies that the theoretical and empirical analysis of algorithmic decisions is, although logical, an implicit evidence. Secondly, this paper acknowledges that given the varying functions and data sources of different hiring software tools, the extent of bias and discriminatory risks differs. Thirdly, given the durational and technical limitations, this research cannot provide an extensive body of research that validates all relevant independent variables.

Part I

Artificial Intelligence and the expansion of inequality

Chapter 1

How do AI algorithms work?

Big Data and Machine Learning are the two pillars that are pivotal to defining AI. Big Data refers to a dataset that is characterised by its massive volume, its high data velocity, its vast variety of data sources and types, and its data validity (Nersessian 2018). Although the sources of these data vary, most data are donated involuntarily by the digital users, such as through social media, website traffic, sensor data and online platforms. This exhibits real-time insights on societal behaviour and social networks (Seely-Gant & Frehill 2015; McFarland & McFarland 2015). Raub (2018) defines Machine Learning as a statistical procedure that automatically generates results from the datasets. This embodies a subset of AI. AI is a family of systems that drives automation to impersonate human intelligence. It improves its own algorithms at each task. Typically, the algorithms generate predictions based on its evaluation on the “interpretable patterns” of the given datasets. Costa *et al.* (2020) succinctly described this process in a diagram as shown in **Figure 1**.

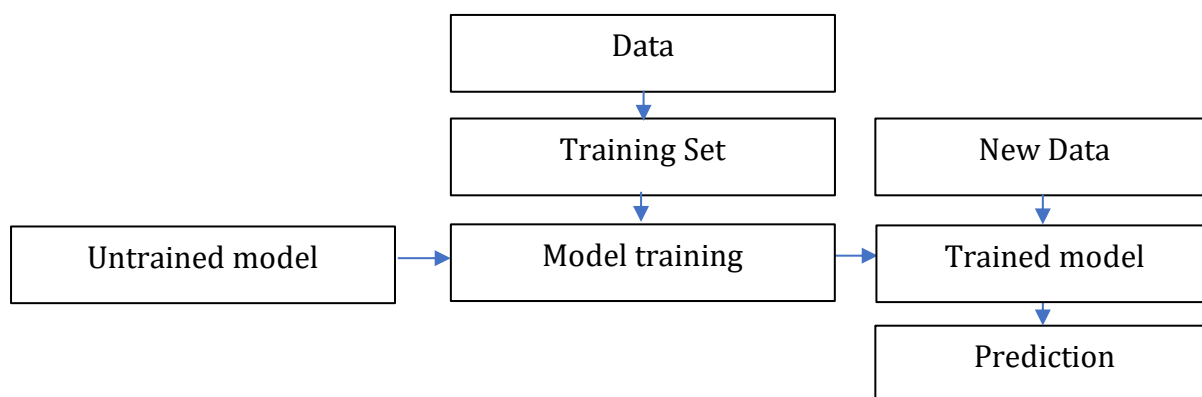


Figure 1: Model Creation Process (Costa *et al.* 2020)

AI is designed to generate outputs using the value of a target variable. Then, its algorithmic model is fed with various types of datasets. The algorithmic model adapts to this dataset to make prediction using new sets of data. This process runs continuously as algorithmic models runs each task.

As indicated, contingent to its function and targeted outputs, different algorithmic models are fed with different types of datasets. In enabling effective predictions, algorithmic

models typically learn *proxy* to represent certain values. One of the practical instances to this is type of dataset is the Big Data. Given its randomised structure, Big Data so often does not contain the desired data values for the designed predictions. Consequently, algorithmic model evaluates the datasets using a *proxy* to represent a characteristic; for instance, sexual orientation, political affiliations, religious convictions and race are often represented by Facebook “likes” (Dattner *et al.* 2019). Given that not all demographic traits are explicitly represented in digital prints, segmenting the digital population requires a certain level of generalisation and representation. Domingos (2012) suggests that algorithms work in amalgamations of three components: representation, evaluation and optimisation. According to Domingos (2012), representation refers to the alteration of ‘language’ of the classifier for the computer to access and learn; evaluation points to the sorting out the invalid classifiers from the valid ones; and optimisation refers to the examining of the highest-scoring classifiers.

Chapter 2

Profit-driven algorithmic judgements: the replica of historical biases

In the past decade, the rise of AI has been geared by marketisation and has given birth to the commoditisation of data. The upsurge of data commoditisation is grounded on the premise that the dynamic movement of information enables efficient algorithmic predictions, unlocking a novel cost-effective decision-making process (Van der Voort *et al.* 2019). Galloway (2017) suggests that there are at least three preconditions that mark the efficiency of data monetisation. Firstly, that data is an economically valuable product, which consists of more than just personal information. Secondly, data are flowing unreservedly from firm to firm, in order to efficiently generate better products for digital users and stakeholders. Lastly, all data are equal, either from the private sector, such as the Big Tech data, including ‘likes’ and search histories; or government data, such as tax, health and income. However, with such liberal data transaction, the notion of power and control cannot remain unchallenged: who has the right to possess and generate data? Who decides to whom the power and the rights belong? Who is the most affected individual and in what way (Van der Voort *et al.* 2019)? This paper contends that, since data and algorithmic predictions have increasingly become a currency for marketisation, generalisation of data becomes a powerful weapon that fortifies the current societal bias.

Algorithmic bias: definitions

First of all, it is important that this paper begins by defining the term ‘algorithmic biases’. In doing so, this paper identifies the three main algorithmic biases as termed by Pauline Kim (2019); namely, the *record error bias*, *intentional bias* and *statistical bias*. *Record error bias* refers to the inaccuracy that occurs due to an incomplete or misrepresentative dataset. This considerably affects the precision of algorithmic prediction and decision-making, especially for the underrepresented groups. *Intentional bias* is the discriminatory action that is intentionally ingrained within the algorithms in order to explicitly filter some groups out of the equation. However, this can also be done for the opposite intention; which is to battle the existing structural inequality. Lastly, Kim introduces the *statistical bias*. This refers to the algorithmic decisions that account demographic traits and combine them with the pre-recorded statistical data of that certain demography. Although this paper acknowledges the importance of all three biases, more attention will be given to the *record error bias* and *statistical bias* in the following sub-sections and chapters. This will refine the scope of discussion as it rules out human intervention as a supplementary variable.

Generalisation and microtargeting: discrimination?

The rise of technology has spurred the highly market-driven digital experience. Digital traffic has become, not only an agency for social interactions, but also a haven of effectual

information-exchange and resourceful data supply. Algorithms categorise the digital population based on demographic traits, including gender, race, neighbourhood and education; and their behaviour, such as preferences, needs, likes and dislikes, as reflected through their digital actions (O’Neil 2016). This is what Shoshana Zuboff (2018) refers to as ‘*surveillance capitalism*’, where data are treated as ‘products’ to be monetised for capitalistic purposes. This behavioural surplus is processed into sets of models to generate predictions and assumptions for the targeted objective. Then, using these assumptions, it predicts what digital content should be directed to which digital population for the highest digital engagement possible. Data are rendered to micro-target a group of segmented demography that is assumed to be very likely to engage in a particular digital information. These contents may include advertisements, digitally shared contents on social media, news content, and so forth. As efficiency is typically the currency for algorithmic success, these predictions are an asset for the companies’ market expansion (Zuboff 2018; O’Neil 2016).

However, these predictions introduce novel risks to maintaining historical inequality. Prescribing future predictions using involuntary data without a controlled environment is a methodology that is exceptionally vulnerable to endogeneity and interpretative errors (Seely-Gant & Frehil 2015). This implies that AI algorithms could potentially produce sheer repetition of recorded human biases, if not worse, an expansion of it.

If the digital population is segmented and microtargeted based on their personal demographic traits, it would imply that AI algorithms apply a discriminatory selection process to determine the subjects of information exposure (Bogen 2019). As an example, Latanya Sweeney, an African American female professor at Harvard University, searched her own name on Google and was directly presented with an advertisement, “Latanya Sweeney, Arrested? 1) Enter name and state, 2) Access full background. Checks instantly.” (Raub 2018). This automated discriminatory segmentation bears problematic implications: not only that all the digitally available information about Sweeney is aggregated and rendered to form a profile for microtargeting, but this set of information is coupled with statistical data to create a correlational point between African American and the risks of crime. The algorithmic prediction, particularly in this case, is grounded on a racial generalisation retrieved from the demographically divided historical data. This implies that, not only is digital information filtered based on one’s digital prints and search history, but also on the patterns of digital users who ‘belong’ to the same group of demographic traits.

This pattern is not limited to race; rather, it encompasses the generalisation of and coupling between gender, income, neighbourhood, ethnicity, religion and nationality, in relation to criminal, professional and educational records, among many others (Maugis 2018; Dennis 2018). Its severity intensifies as algorithms are designed as a means for profit-oriented advertisements and predictive judgements. In 2013, a for-profit college in the United States

was found to intentionally advertise its institution to ‘vulnerable’ digital users for the sake of effective marketing (O’Neil 2016). They use the words “low self-esteem”, “isolated”, “stuck”, “unable to plan well for the future”, and “impatient” as their microtargeting keywords. Unsurprisingly, there appears to be a pattern of personal demographic traits to these descriptions, including its consequential linkage to certain neighbourhoods and race. It is apparent that, given its proneness in making spurious auto-correlational errors, algorithms pose the risk of replicating institutional and historical biases, prowling on data points that could affect one’s prospective opportunities based on their demography. As this paper will explain in the following chapters, this includes the microtargeting of job vacancies.

Furthermore, algorithms are so often perceived as a ‘black box’ (Faragher 2019). Its overly complex and impenetrable nature makes algorithmic judgements inexplicable, even by its own programmers. Primarily, given that unveiling the ‘black box’ is rather impossible, it is almost impractical to detect any faulty algorithmic calculations or defective predictions. Consequently, the correlational points can be blindly drawn between demography and any other attributes to generate conclusions without being thoroughly justified or even pre-examined. Hence, it concedes a fallibility in the ensuring a fair accountability mechanism (O’Neil 2016). It would be particularly abstruse to hold any algorithmic discriminatory conclusion liable for its decision.

The problem of misrepresentation

As outlined, algorithmic models are foundational to galvanising algorithmic biases. Aside from the undetectable algorithmic inputs and outputs, the issue of ‘representation’ is critical to social and ethical considerations. The subjects that are excluded from the datasets should be equally and carefully considered as those who are represented. However, without conscious deliberations, most datasets, including Big Data, are prone to exemplifying unstructured data; in which case, the datasets are obtained from solely observational means rather than extracted from a controlled environment using rigorous methodology (McFarland & McFarland 2015). This bears an extensive leeway for noise, projecting multiplied, missing and inaccurate values, and consequently, can cause a serious problem of misrepresentation.

In particularising the issue of data representation, this paper acknowledges the significance of dissecting the digital population and considering the digital divide. In 2018, only 69 percent of the total Australian population are digitally active users, with only 53 percent to be the global average (Kemp 2018). A report by Barraket *et al.* (2019) on Measuring Australia’s Digital Divide eloquently describes the current digital inclusion gaps in Australia using the Australian Digital Inclusion Index (ADII). The digital divide in Australia is significantly correlated with socio-economic backgrounds, including income, education and employment. Individuals within the low-income households are considerably excluded from the digital realm. There is also a significant digital inclusion gap for the Australian minorities, including people with

disabilities and Indigenous communities. Furthermore, the gap between digital population and the general population is much larger in the rural and country areas in comparison to the metropolitan citizens. Women are reported to score 1.8 points less than men in ADII, which is relatively narrow in contrast to previous years.

The statistics evidenced above logically invoke the question: what happens to those who are *not* represented in the data? This paper proposes two theoretical presuppositions. Firstly, when algorithms take part in a decision-making process, it is evident that the underrepresented individuals are out of the equation. Secondly, as AI improves its algorithms at task, the algorithmic model lessens its sensitivity in consideration of the underrepresented groups as it adapts to the absence of these groups in improvement of the model. This implies that the algorithms would form a model that work in great favour for the represented groups but would run much more ineffectively for the others. For instance, a research conducted to find the contributing factors to marital divorce is majorly based on traditional marriages between a female and a male (Seely-Gant & Frehill 2015). The missing data on other contributing factors such as marriage with more than one partners or between LGBTQ+ communities would affect how the algorithmic models run in finding this causation point.

This points the case towards the second concern of misrepresentation, which is the misrepresentation of *proxy* (Dalenberg 2018). As a hypothetical instance, if the algorithmic models capture a correlational point between the success of law graduates, their GPA and to the fact that the majority of them play tennis, then the algorithms would generate a model that grounds its success prediction based on the high-scoring GPA *and* whether or not they play tennis. This is often referred to as the correlational or statistical bias; which means that the attributes captured as correlational do not always mean it forms a causational relationship. However, the case of Latanya Sweeney evidently shows that algorithms are prone to this faulty assessment. It captures an unnecessary correlation between an African American to the likelihood of needing a lawyer. This poses a significant question mark to the definition of 'fairness' in algorithmic decision-making: what data values are to be considered as valid determinants in making a prediction for a targeted output? Suppose that the hypothetical law graduates' case is applied in the case of the job market, then the algorithmic would be largely affected by whether or not the candidate plays tennis, rather than professional contributions.

The algorithms' susceptibility towards correlational bias also triggers the misrepresentation of logic. One of the agonies of AI is its ability to assess a massive volume of datasets, including Big Data; in other words, the quantity of data. However, as indicated algorithms predominantly expose the networked relationships between datasets and between values, but algorithms cannot explain *why* they are relational (Chauveron 2016). Returning to the case of the law graduates, it is apparent that there is a link between the success metrics and tennis-playing, but the data do not explain why the connection exists. It is severely

problematic when these typical algorithmic predictions are applied using randomised datasets. Generally, as large samples raise the sensitivity of a single change, the possibility of inaccuracies also expands (McFarland & McFarland 2015). By recognising the prospect of misrepresentation, extracting algorithmic judgements without careful considerations on the algorithmic models and the representation of the data could be a dangerous method (Athey 2017). Datasets may travel through sequences of information supply chain and, therefore, are possibly refurbished, reused, reinterpreted and reprocessed. Without adapting to socio-economic, environmental, cultural and other contexts, algorithms may fail to recognise the core meaning of the data itself (Schintler & Kulkarni 2014, p. 344).

Statistical bias paving the way for systemic bias

These questionable inferences highlight the same human problem, which is the subconscious bias (Cohen 2019). The case of data misrepresentation is not limited to the use of Big Data. In fact, algorithms are prone to making correlational errors and in creating unintended discriminatory decisions even with selected datasets. Without rigorous rubrics of algorithmic evaluation, discriminatory outcomes can be severe (Seely-Gant & Sehill 2015). Suppose that an algorithmic model is designed to predict the potential success of a student candidate for a college admission. The profile samples for the measure of ‘success’ remains ambiguous. What determinants are used to construct the idea of a ‘successful’ candidate? Whose profile from which these determinants are derived? If the model is based on the college’s student retention, what does the college demographic diversity look like? Using these, would algorithms discriminate against other demographic traits because of their demography rather than their merits? These questions are pivotal to minimising bias in the algorithmic formation. Nonetheless, these decisions are rarely inquired and are largely determined by a handful of tech actors.

The key literature has identified this issue to be majorly influenced by the pertaining lack of diversity within the tech industry (Raub 2018). In 2016, three major tech companies in Silicon Valley did not employ African-descent employees and ten others did not have any female African American employees (Rangarajan 2018). In 2017, the Centre for Investigative Reporting sent inquiries to 211 technology companies in Silicon Valley to report their employee demographics to the federal government (Cao 2017). Despite the absence of legal obligation for a response, 22 companies, including the giant techs, such as Facebook, LinkedIn, Apple and Google (FLAG companies), responded to the inquiry. It is found that, not only are these tech companies managed by males, but most of the executive positions are occupied by *Caucasian* males. In fact, Asian workers over-represent the minorities across different companies. African Americans and Latinos in aggregate only account for single digits. Despite the United States Government’s attempt to resolve the lack of diversity, it appears that tech companies find great difficulty to attract talented female candidates (Raub 2018).

The numbers have revealed the endogeneity within the tech businesses. Raub (2018) identified the severe challenge of acknowledging the value of diversity and aiming to minimise inequality in an environment which does not. On the contrary, the tech giants often detect inequality and beneficially condone it (O’Neil 2016). The repeated class imbalance, racial segregation, gender inequality, along with other forms of ethical concerns, are often used to escort the tech businesses to the forefront of the capitalistic market.

Chapter 3

Algorithmic biases in determining hiring decisions

Corporate recruitment has been identified as one of the key components that determine the success of an organisation (Kulkarni & Che 2017). Consequently, recruitment has transformed into a broader concept of “Talent Management” (TM), which is an umbrella term encompassing a wide range of data collection for corporate human resources management. This includes the cycle of recruitment, performance evaluation and incentive execution (Schweyer 2010). This transformation is spurred by the economic principle which perceives talent acquisition as a means to minimise cost for maximum corporate benefits (Kulkarni & Che 2017). As a consequence, recruitment does not only entail talent retention and acquisition, but it forms a system that calculates the corporate loss and gain for every employee replacement.

The rising demand for efficient recruitment escalates the use of hiring algorithms as a recruitment tool as the most efficient automated applications tracker (Van Esch *et al.* 2019). This particular trend is triggered by four factors: the dynamic imbalance between job demand and supply; the changes in business management which so often requires costly technological outsourcing; communicative and technological advancements; and the upsurge of data analytics methodologies which enable ‘success’ predictions using detailed calculations in a candidate assessment (Kulkarni & Che 2017). In adapting to these challenges, algorithms combine behavioural factors, such as typing rhythms and voice patterns, and physiological factors, such as iris movements and micro-expressions, among others, as a part of its broader decision-making process (Van Esch *et al.* 2019). It progressively adapts to different tasks to amplify its accuracy and effectiveness for prediction-making.

As recruitment involves a culmination of incremental decisions, hiring algorithms interfere in different stages with varying functions using different processes (Bogen 2019). Kulkarni and Che (2017) categorise hiring software tools into three main distinctive functions. The first model is the *Job Aggregator Software* (JAS), which compiles different job vacancies from a number of different websites and posts them into one website with the aim to attract maximum traffic of candidature. The second model aims to evaluate the suitability of candidates using predetermined criteria and measures or a set of tests conducted by the software. This is referred to as *Candidate Assessment Software* (CAS), with examples such as TestDome, eSkills and so forth. Related to the second, *Applicant Tracking Software* (ATS) refers to the bureaucratic aspect of recruitment, which typically executes the following tasks: create and advertise job vacancies, collect resumés, shortlist applicants in accordance to a set of criteria, schedule interviews for shortlisted applicants, perform interviews, send offers to selected candidates and assist the onboarding procedure. Ceipal TalentHire, Jobvite and Bullhorn are examples of this software. This tool typically enables companies to comply with

the Equal Employment Opportunity Commission regulations by keeping resumés in one storage. This paper, however, will focus more on CAS and ATS, which will be explained further later in the chapter.

Automated hiring has proven to be an impactful innovation in major companies. Nevertheless, the main issue of discrimination remains a significant concern. Firstly, if the algorithms can capture, memorise and predict characteristics such as physical features, gender, health, race, sexual orientation and age, it is crucial to question whether the algorithms use these attributes as a determining factor to candidates' selection (Van Esch *et al.* 2019). Following this, it is also vital to question whether the meritocratic measures to the selection process are being tied to these characteristics to make generalisation of the 'success' prediction (Roy 2017). The following subsections will expand on the pertaining biases that remain inherently ingrained within the wholeness of recruitment algorithms: in job advertisements and candidate selections.

Jobs ads: maximum engagement through microtargeting

Job advertisement is principally designed for two particular purposes: firstly, the job vacancies should be directed to the most suitable candidates, and secondly, the suitable candidates should be directed with the right and engaging advertising contents (Dalenberg 2018). The objective of this is to attain the most 'meaningful' engagement with candidates, measured by the number of advertisements 'clicks' as the determining currency of success (Bogen 2019). Based on these suggestions, algorithms appear to be a direct response to this objective. It promises well-curated platforms that generate highly superficial categorisation of subjects and predictions on the likelihood of engagement using a variety of determinants, such as location, age, gender language, education, job, income, relationship status, and more (Dalenberg 2018). These data are extracted from certain datasets, most typically Big Data, which consist of random information that is then processed to generate patterns and make assumptions for predictive purposes. As mentioned in the previous chapter, however, these concluding patterns are not of controlled order; on the contrary, these patterns are generated using statistical correlations, which most of the time do not infer any logical causation. Furthermore, in order to pipeline a maximum amount of engagements, algorithms memorise digital actions and behaviour, including the dynamic interactions between employers and candidates (Bogen 2019). It records the characterising attributes from the interactive patterns to improve its accuracy for future microtargeting. This means that, if the algorithms perceive that most of the interactions are undertaken by Caucasian male candidates, the algorithms will tend to target Caucasian men for similar job advertisements in the future. These algorithms continuously learn the dynamic patterns of engagements and interactions until the automated system reaches the intended audience with the most engagement (Dalenberg 2018). As a consequence, AI replicates the illogically biased causations as the machine refines its own calculations.

This raises the concerns for discriminatory outcomes. Even without the direct human interference, the typical cases involving hiring algorithms can replicate the pertaining societal biases (Bogen 2019). Almost in all cases where job advertisements are controlled by algorithms, the individuals who are categorised within the protected groups remain those who are most likely to be negatively affected (Lambrecht & Tucker 2020). Bogen (2019) and colleagues conducted a study in which they observe the microtargeting of job advertisements on Facebook. It is found that supermarket cashier job vacancies are presented to a population of 85 percent women whilst taxi driver job vacancies are shown to an audience of 75 percent individuals of an African race. In aggregate, job ads for high-salary executive positions are presented to men six times more often than women (Dalenberg 2018).

Lambrecht and Tucker (2020) conducted a study to investigate the microtargeting of STEM (Science, Technology, Engineering and Math) job vacancies. With its intention in responding to the policymakers' concerns about the underrepresentation of women within the field, careful consideration was put into gender-neutral targeting. The advertisement was tested in 191 countries to individuals in their most productive years. With this controlled environment, the outcome shows that the advertisement is shown to 20 percent more male than female audience. Lambrecht and Tucker (2020), however, suggest that this is not based on the likelihood of ads engagements; in fact, it is found that women were more likely to engage with the ads in comparison to men. However, it is suggested that as a consequence of this, women are a high-quality demographic and therefore are a more high-priced population as the object of advertisements. In adjusting to the cost effectiveness of these advertisements, therefore, they are shown, not according to gender balance, but cost and profit balance. It is, therefore, concluded that, with the capitalistic principle, algorithms are prone to generate discriminating outcomes even in relatively and intentionally the most 'neutral' setting (Lambrecht & Tucker 2020).

As evidenced, job advertisement microtargeting algorithms are likely to adopt biased judgements, which may entail: the seemingly prejudiced behaviour in contrast to the real consumer behaviour; the discriminatory patterns against certain groups across different countries; and the economic consideration as the main priority for job advertising. This issue cannot remain unchallenged. The absence of information on opportunities is a highly significant hindrance to an individual's right for fair opportunity (Kim, as cited by Bogen 2019). It reflects the current flaw of the praxis of equal opportunity and equal access for all, especially for women (Dalenberg 2018).

Job sorting: translating past history into future outcomes

The adoption of hiring algorithms was initially grounded on the premises that, firstly, the absence of human intervention gives rise to impartiality; and secondly, that technology

enables efficiency and accuracy in sorting a massive volume of applications with minimum cost and for maximum benefit of the company (Preuss 2017; Kulkarni & Che 2017). Theoretically, AI *should* be able to create an optimum amalgam of excellent candidates on the basis of pure meritocracy. As indicated, this paper seeks to invalidate this premise. In contract, it argues that the market-driven algorithms are radically prone to replicate societal biases. In making this case, this subsection will explain the theoretical logic on the proneness of the algorithms to making correlational and statistical bias in the prediction of a suitable candidate.

It is, however, important that this paper first expound on the variety of commonly adopted hiring tools and the different functions of hiring algorithms. Commonly, different tools would perform identical principles in running a set of recruitment tasks. It would typically source and screen data, from either direct input (including tests, assessments, video interviews and submitted resumés), or from automated input (usually from Big Data). These data would then be used to grade and rank candidates' suitability for the role based on weighted components, such as assessment metrics, qualifications, test scoring, keywords and so forth (Roy 2017). However, different software tools perform using different approaches in accordance to its designated objectives. Although there are various categorisations of automated hiring software, this paper focuses on recruitment tools that function as *Candidate Assessment Software* (CAS) and *Applicant Tracking Software* (ATS). CAS tools typically present personality tests, case studies, logic tests, automated interviews and other assessment metrics and assess the candidates' test performance against the predetermined set of criteria. Generally, the tests are classified into three distinctive categories: intelligence, personality and mental or clinical health tests (Dattner *et al.* 2019). One of the examples to CAS tools is *HireVue*. This software assesses candidates based on virtual interviews, among many others. The algorithmic model ranks candidates' suitability and success predictions using voice and facial recognition (Raub 2018). Software tools such as this can evaluate the candidates' choice of words, for instance, to predict the levels of empathy, hospitality, and other representative characteristics, and weigh these attributes in accordance to the company's culture (Alsever 2017). Similar idea applies to logic tests, where candidates are asked to perform assessments to be ranked for their suitability and success metrics (see, similar tools such as *ARYA*, *CEIPA* and *HackerRank*).

ATS, on the other hand, performs resumé parsing, *Customer Relationship Management* (CRM), background and social screening, candidates sourcing, and success prediction. It is to note, however, that not all of ATS tools would include all functions. Some can only perform one specific function whilst others may be able to perform all functions, and a few other ones may include both CAS and ATS operations. One of the most prominent examples to this is *Gild*. *Gild* performs beyond resumé parsing and extends its rendition from the Big Data (O'Neil 2016). It implies that the assessment of candidates' suitability is not only based on qualifications or meritocracy, but it also rates applicants based on their 'social capital',

generated from their digital activities. An identical approach applies to other tools, especially those involving background screening from Big Data, such as *Ascendify*, *BambooHR* and others. Major companies with a massive amount of applications are reliant on these tools for their hiring process (O’Neil 2016). In fact, seventy-two per cent of resumés are never seen by hiring officers at all (O’Neil 2016). It is to note that, given the numerous ways in which these different tools source datasets and perform tasks, the degree of risks may vary. Nevertheless, this paper suggests that, despite its diverse models, algorithmic hiring decisions are highly sensitive to making biased judgements.

Firstly, the limitation of the datasets acts as a major factor to the algorithms’ proneness to discrimination as it builds the scope of benchmarking (Costa *et al.* 2020). If the algorithms are fed with limited datasets to generate decisions, the algorithms will limit its ‘success’ metrics based on this particular set of data (Costa *et al.* 2020). For instance, an experiment was done using a facial recognition software to see its accuracy when applied to a variety of racial differences. Due to a large amount of inputs from Caucasian participants, the software performs higher accuracy when observing Caucasians, and lower accuracy towards the races of minority (Costa *et al.* 2020). This validates the presupposition indicated in Chapter Two that data representation is highly crucial in determining an accurate prediction. In the same manner, considering excluded demography within a dataset is equally significant as those who are represented.

Secondly, as evidenced in the previous chapters, correlational bias remains a major concern in algorithms, including those designed for hiring management (Kim 2019). One of the prominent reasons for this is the use of a *proxy* to represent a candidate and the parameter of suitability. As an example, algorithms make a correlational relationship between ‘creativity’ and an individual’s length of employment within the respective job (O’Neil 2016). Algorithms also make associations between higher levels of ‘inquisitiveness’ and an individual’s likeliness to find other opportunities. When these correlations are conjoined with demographic traits as a *proxy*, such as neighbourhood, race or gender, it poses severe problems as it could infiltrate the entire discriminative corporate culture (O’Neil 2016). When these *proxies* are embedded within the algorithmic judgement of employment measures of success, it can reflect, if not, expand societal bias. In background screening, for instance, given that the algorithms seek historical records using the Big Data, the decisions will be more sensitive in making racial biases, especially with regards to various discriminatory records across the history (O’Neil 2016; Raub 2018). In the 1980s, a medical school in the United Kingdom, St. George’s Hospital, used a hiring program for its admission decision (Raub 2018). It was found that, instead of introducing new biases, the algorithms captured the already existing biases as its target variables, and therefore mimicking these discriminatory decisions as its own evaluation patterns. Consequently, every year, the software rejected approximately 60 applicants on the basis of race and gender.

These instances should be a call for a re-examination towards the legitimacy of the 'suitability' parameters. This paper has identified the various ways in which algorithmic bias can occur since the beginning of the suitability assessment process. In other words, if each candidate is to be assessed using the same set of criteria, where will this model of 'ideal' candidate come from? How much of the algorithmic judgement would be determined by 'objective' measures in comparison to the correlational variables? The inputs of these criteria can be severely problematic if not considered rigorously (O'Neil 2016). It applies in a similar manner both in CAS and ATS tools. Firstly, the definition of an 'ideal' employee can be ambiguous. Ideally, when recruiters decide a certain set of criteria, they will settle on measurable outcomes, such as GPA results, longer employment retention, and so forth (Raub 2018). However, the acknowledgement of the persisting subjective decisions and heuristic biases in human decisions, the case for hiring algorithms speaks for itself (Raub 2018). Nevertheless, this paper argues that the adoption of hiring algorithms does not in any way eliminate these biases on the basis of two reasons. Firstly, algorithms are generally trained to record and memorise past decisions and *learn* from them (O'Neil 2016). This implies that the algorithms memorise the patterns of previous decisions and replicate these patterns in algorithmic decisions. Secondly, in the cases where algorithmic judgement becomes conclusive in the hiring decisions, there is a risk of compromising the objective measures. As the algorithms adapt to the preceding human decisions, it can transform their assessment rubrics according to subjective considerations for the sake of effective and efficient hiring predictions (Raub 2018). This is expected to generate outputs with the most cost efficiency as it predicts candidates' suitability based on dynamic determinants, regardless of its compromise towards impartial judgements (Kulkarni & Che 2017).

The suggested contentions direct the case towards the impact of hiring algorithms on gender equality, especially with regards to gender-based representation in the labour force. If the male population has been in the workforce longer than female, how would that translate into algorithms in determining the benchmark of success? Correspondingly, if women who are currently in the workforce are underrepresented in full-time labour, and occupy most of the casual, part-time and domestic work, how would this impact the future of other women (Australian Government 2020)?

Part II

Algorithmic bias against women: the mirror of societal discrimination

Chapter 4

Correlational bias in hiring: women's future in the hands of men

The idea that hiring algorithms select the most suitable candidates with minimum cost and maximum corporate benefits becomes the engine of meritocracy (Kunovich & Slomczynski 2007). However, this paper argues that, with algorithms' sensitivity to correlational bias and its self-evolving nature, the meritocratic principle bears the risk of becoming a subjective idea.

As an example, in 2014, Amazon generated hiring algorithms to parse resumés and predict the suitability of applicants. The algorithms are then with the existing data from their own workforce over the past 10 years (Costa *et al.* 2020). In 2018, it was found that Amazon's hiring algorithms discriminate against female applicants (Dastin 2020; O'Neil 2016; Bogen 2019). However, this discrimination was not intentional; rather, it reflects the very pitfall of algorithmic bias itself: it perpetuates the existing inequality and societal bias (O'Neil 2016; Costa *et al.* 2020). As the majority of Amazon's employees consist of white male employees, their hiring algorithms capture this pattern as a determining factor of success, and therefore, develop a bias against the female candidates (Costa *et al.* 2020; Faragher 2019). This includes discriminating applicants with keywords such as "all-women's college", "female" and other words that indicate that the candidates are female, and rank these applicants lower in score (Costa *et al.* 2020; Faragher 2019). Amazon shut down this system soon after this algorithmic bias was found (Dastin 2020).

The Amazon case has presented that the correlational and statistical bias bears a significant discriminative against the female population. Faragher (2019) also suggests that, given the homogenous nature of tech industry's demography, unintentional discrimination may be easily missed or overlooked. This chapter, then, will provide an in-depth outlook towards its hypothetical implications, especially with regards to the existing imbalance in gender representation. The first section will make a logical inference to algorithmic biases as a consequence of a limited training dataset. It will present the existing underrepresentation of women in data and what this implies to statistical discrimination. Following this, the second section will explain the implication of correlational biases in the suitability determinants used to assess female candidates. It will include the discussion surrounding the measures of 'suitable candidate' and the impact of gendered use of languages and keywords.

The underrepresentation of women in labour and data

As evidenced by the Amazon case, this paper contends that, if the statistical and correlation bias against women can be caused by the limitation of the dataset, then the representation of women in labour data should be of a critical concern. In making this case, this paper will first investigate the existing labour market based on gender and the representation of women in workforce data, globally and specifically in Australia. It will, then, assess its inferences by drawing on existing research and theoretical analyses.

Although gender equality in the labour market has improved over the last few decades, the global workforce remains largely occupied by men (Ortiz-Ospina & Tzvetkova 2017; International Labour Organisation (ILO) 2019). In 2019, whilst 74.7 percent of all men participate in the workforce, less than half of the women do (The World Bank Databank 2019). In 2018, the European Union reported that only 48.3 percent women are employed in formal economic labour (Eurostat Database 2019); this is similar in the United States with 57.1 percent of women participating in the labour force compared to 68.1 percent of men (U.S. Bureau of Labor Statistics 2018). All over the world, women dominate the majority of labour works that are not formally documented as ‘economic activities’ (ILO 2019).² In 2019, it is reported that out of the 500 leading executives in the most prominent firms around the world, under 7 percent are women (UN Women 2020). In developing countries, the occupational segregation between men and women remains high, despite the increasing educational quality (UN Women 2020).

Similarly, the outlook of Australian workforce is as severe. The Australian Government (2019) reported that in 2019, 5.2 percent of Australian women are unemployed, in comparison to Australian men at 4.9 percent. Female underemployment, which refers to those in the labour force who aim and are available for more working hours, is at 10.4 percent, whilst male underemployment rate is at 6.2 percent. Lastly, amongst the entire Australian part-time employment, 68.7 percent are occupied by women. In Australia, men have a higher job retention rate and transition rate into employment (Melbourne Institute 2019). This is in large part a consequence of motherhood. On average, an estimated rate of 25 percent of unemployed men aged 18 to 64 enter employment, whilst only approximately 5 percent of men leave employment. On the other hand, 20 percent unemployed women of the same age transition into employment at any given year, whilst 10 percent of employed women transition out. Women are also more likely to transition out of full-time employment, to either unemployment or to part-time workforce, in comparison to men.

² ILO defines informal ‘economic activities’ as work within the workforce owned by households, including informal self-own enterprises and undocumented business or business owners. ‘Undocumented’ is vital to this paper.

The underrepresentation of women in the labour force data, then, should be seriously considered in the algorithmic data training. This paper, thus, contends that, firstly, if the algorithms are fed with the existing internal workforce data, algorithms can generate correlational bias between meritocratic principles and gender, and therefore, would discriminate against women. Secondly, if the algorithms are trained using Big Data in conjunction with the submitted resumés, the validity of data becomes questionable, especially when extended to a candidate's social data. As a hypothetical example, if the algorithms capture a candidate as being married, would they, then, assume that the candidate would be more prone to exiting the workforce because of motherhood? The principle of meritocracy in conjunction with the correlational bias, then, bears a very significant risk to box women out of the workforce. Thirdly, as an implication of the two points above, if algorithms assume that the model of a successful candidate is male, then the determinants of 'suitability' itself need to be reconsidered.

Gendered roles, gendered algorithmic judgements

It is evident that women's position in the society has shaped the way in which women are socially perceived (Sczesny *et al.* 2016). This translates into androcentric biases as a norm in describing both men and women (Hegarty & Buechel 2006). Androcentric pronouns and references are more generally used to describe both sexes in data and documentations. Men are perceived to be a more typical member and the universal standard of 'human' commonly cited. On the contrary, in categories where women are overrepresented, women are perceived to be the typical member of that particular group; for instance, when it comes to describing 'parenting' (Hegarty & Buechel 2006). These inaccuracies and generalisation of data become problematic when trained into algorithms as they may generate results that are equally inaccurate or worse, expand the inaccuracies by blindly making decisions based on these data.

Key literature on the fairness of hiring often mention the impact of language in gender discrimination, from the use of masculine and feminine words in job advertisements to the use of certain words in describing oneself in job applications (Stout & Dasgupta 2011; Gaucher, Friesen & Kay 2011; Horvart & Sczesny 2016). Gaucher *et al.* (2011) have found that there are gender differences in the use of language and psychological traits between men and women. It has been reported that women would make reference to more communal words and utilise social and expressive words in comparison to men (Gaucher *et al.* 2011). Not only that men and women tend to use different adjectives to describe themselves, they also use different adjectives to describe other people according to gender. For instance, Schmader, Whitehead & Wysocki (2007, cited by Gaucher *et al.* 2011) have found that letters of recommendations that describe men contain words that describe prominence, such as 'outstanding' or 'unique'. On the contrary, letters that describe women suggest words that contain more social and less agentic connotations, such as 'warm' and 'collaborative' (Sczesny *et al.* 2016). As evidenced by the Amazon case, the use of keywords that describe gender, such as 'women's chess club

captain' or 'women's college' is highly critical (Dastin 2020). However, it is also evident that the problem does not stop at keywords, but it expands to the use of feminine and masculine languages. This becomes a significant contribution to algorithmic bias as algorithms learn from historical data and protected groups, including women, have been underrepresented in the labour force, especially in executive positions (Australian Government 2019). It bears the risk of filtering out feminine-worded resumés (Faragher 2019). In fact, numerous algorithmic hiring tools have boxed out CVs that entail feminine words, such as 'collaborative' or 'supportive', in comparison to those that include more masculine words, such as 'execute' or 'competitive' (Faragher 2019).

Chapter 5

Identical but gendered: hiring algorithms in assessing candidates

This paper intends to test the underlying hypothesis and to validate its theoretical inferences. For this purpose, this study has designed an experiment in four stages for four different research questions. All four stages will be directed to manipulate the hiring algorithms using fake resumés that are altered using gendered variables. These variables will change according to the research sub-questions. In minimising confounding variables for all the four stages, a set of criteria apply, such as: 1) targeted jobs are to be the most gender-balanced in Australia (which has been identified to be ‘accounting jobs’ (Tilley 2018)); 2) fake resumés are to be made based on a profile of a 35-year old person in their mid-career; 3) the targeted job vacancies are to require 7-10 years of experience; 4) fake resumés are to indicate excellent career progression at well-known companies; 5) the resumés have to be submitted to at least 40 job openings for significant results; and 6) job postings have to only be full-time jobs.

Stage One: James and Jane study

Methodology

The first stage of the experiment will aim to answer a broader question, *do hiring algorithms discriminate against women in selecting candidates?* In answering this, this paper hypothesises that hiring algorithms prefer male candidates instead of female, despite the identical attributes. In validating this, this experiment will minimise as much variable as possible in order to ensure that the algorithms make decisions only based on gender. The experiment begins by sending through two exact same resumés with two different names, which clearly identify different genders: male and female. These two names are *James Sachs* and *Jane Seaman*. The two resumés are submitted to 40 *Corporate Accountant* or *Head Accountant* jobs in a span of one week. The measured variables would be the number of call-backs and emails from employers, either positive (proceeding into the next hiring stage) or negative (application rejections), comparing the numbers between James Sachs and Jane Seaman.

Result and discussion

The observation shows that Jane Seaman received 16 rejection emails, whilst James Sachs received 12 rejection emails from the same companies for the same roles. However, the 4 roles from 4 companies have not contacted James Sachs in any way and have not identified specific reasons for the different outcomes. On the other hand, Jane Seaman has not received any call-backs nor email responses indicating company’s interest in the resumés to proceed into the next hiring stage. It is, however, found that James Sachs received 15% more positive

responses than Jane Seaman. Even with the limited data samples, the results evidently demonstrate that there are underlying biases in the hiring algorithms. The resumés have the exact same components with the exact same structures with only the names of the candidates being different. This implies that the hypothesis of experiment stage one is proven to be true.

Table 1. Results of the first stage experiment

<i>Subjects</i>	<i>No. of job applied</i>	<i>No. of positive calls</i>	<i>No. of positive emails</i>	<i>No. of rejection emails</i>	<i>Total no. of responses</i>	<i>% of positive responses*</i>
<i>James Sachs</i>	40	6	1	12	19	36.8%
<i>Jane Seaman</i>	40	0	0	16	16	0%

*in comparison to all responses (not total job applications).

Stage Two: online job profiling study

Methodology

Stage Two aims to answer the question: *do hiring algorithms in online candidates profiling discriminate against women?* By online candidates profiling, this paper refers to the candidate's profile matching sites such as *LinkedIn*, *Seek.com.au*, and so forth. The hypothesis of this experiment is that, similar to the resumé parsing algorithms, the profile matching algorithms would also bear gender biases. This means that prospective employers or hiring agencies would find more male candidates than female candidates on their screens as the top successful candidates, even when the profile contents between the two are of the exact same resumé. To confirm this hypothesis, this experiment will attempt to manipulate an online candidate profiling platform by filling in the profiles of James Sachs and Jane Seaman using the same resumés used in Stage One. The measure of outputs will be the number of resumé downloads and profile views. This observation is done in a span of four weeks.

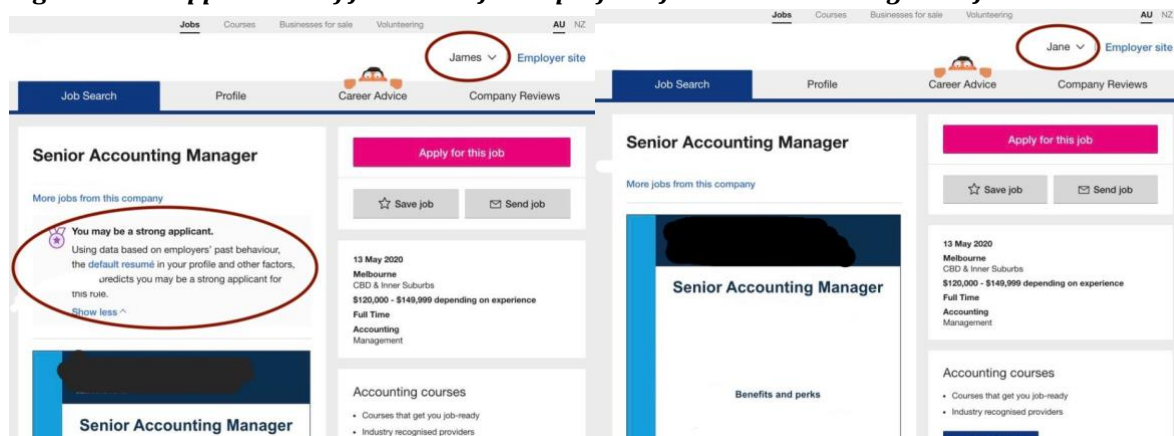
Result and discussion

The result demonstrates a significant difference in the number of resumé downloads between James Sachs and Jane Seaman. Whilst Jane Seaman only receives 1 profile view from a prospective hirer, James Sachs receives 16 profile views. Furthermore, whilst Jane Seaman does not receive any resumé download from any prospective hirer, James Sachs receives 9 resumé downloads from different prospective hirers. Additionally, every relevant job posting presents a badge on James' account indicating that he is a 'strong applicant', whilst none appears on Jane's, despite the identical profile, aside from names. This paper contends that this has validated the hypothesis that, just like resumé parsing algorithms, profile matching algorithms also bear gender biases. The matching function appears to work better with male candidates than female.

Table 2. Results of the second stage experiment

Subjects	No. of profile views	No. of resumé downloads
<i>James Sachs</i>	6	1
<i>Jane Seaman</i>	0	0

Figure 1. The appearance of James and Jane's profiles: James with the badge and Jane without



Stage Three: motherhood and gender study

Methodology

The next stage explores the questions, *do hiring algorithms discriminate against women who are in or have been through motherhood? Does parental leave make a difference in job hiring?* This paper hypothesises that the algorithms will predict less successful outcomes for women who are in or have been through motherhood in comparison to male candidates who have taken the same amount of parental leave gap. In testing this hypothesis, the experiment will use four resumé that are altered according to two variables: gender and parental leave. The first two resumé, James Sachs and Jane Seaman, will be of the exact same resumé used in Stage One. Similarly, the other two resumé will have male and female names, James Smalls and Jane Sparks. James Smalls and Jane Sparks' resumé are the exact same resumé following the previous resumé with a three-year gap from 2017 to 2020, clearly indicated due to parental leave. These four resumé are submitted to 40 different job roles. This paper presumes that the resumé without gap years will be far preferred than the two resumé with parental leave; however, this paper presumes that male candidates with gap years will be preferred than female with gap years. If this is proven true, then, motherhood is considered to be a disadvantage by the hiring algorithms; and therefore, with everything being equal, a disadvantage for females in pursuing future workforce.

Result and discussion

The result shows that out of all job roles, James Smalls and Jane Sparks (candidates with parental leave) receive 10% more direct rejection emails than the others. However, there is no indication that there is any different treatment between male and female in this matter. James Sachs (male candidate without gap years) leads by 10% in comparison to everyone else in terms of positive call-backs. It is, therefore, evidenced, that parental leave between two genders is not an impactful variable to job applications; however, it remains true that the existence of parental leave is a significant variable when compared to candidates without gap years.

Table 3. Results of the third stage experiment

<i>Subjects</i>	<i>No. of job applied</i>	<i>No. of positive calls</i>	<i>No. of positive emails</i>	<i>No. of rejection emails</i>	<i>Total no. of responses</i>	<i>% of positive responses*</i>
<i>James Sachs</i>	40	6	1	12	19	36.8%
<i>Jane Seaman</i>	40	0	0	16	16	0%
<i>James Smalls</i>	40	0	0	18	18	0%
<i>Jane Sparks</i>	40	0	0	18	18	0%

*in comparison to all responses (not total job applications).

Stage Four: gendered keywords study

Methodology

The last stage of the experiment investigates the question, *do hiring algorithms discriminate women more when resumés are altered using feminine or masculine keywords?* As indicated in the previous chapter, one of the key variables that create gender bias in hiring algorithms is candidates' use of words. It has been indicated previously that there are adjectives and verbs that are considered to be used more by female candidates, such as 'collaborating', 'supporting', 'facilitating', and more; and those that are used more by male candidates, such as 'directing', 'leading', and more. In this experiment, therefore, this paper hypothesises that the algorithms will be in favour of resumés that utilise masculine words in comparison to feminine, even when the experiences and skills are exactly the same. In testing this hypothesis, this experiment will use James Sachs and Jane Seaman resumés as indicated in Stage One; however, the keywords will be altered according to feminine and masculine language.

Result and discussion

The observation shows that Jane Seaman received 17 rejection emails, whilst James Sachs received 5 rejection emails from the same companies for the same roles. On the other hand, Jane Seaman has only received one email response indicating the company's interest in the résumé. James Sachs received 17.5% more positive responses than Jane Seaman. Again, the results evidently demonstrate that algorithmic biases occur more with gendered keywords than without.

Table 4. Results of the fourth stage experiment

<i>Subjects</i>	<i>No. of job applied</i>	<i>No. of positive calls</i>	<i>No. of positive emails</i>	<i>No. of rejection emails</i>	<i>Total no. of responses</i>	<i>% of positive responses*</i>
<i>James Sachs</i>	40	3	4	5	12	58.3
<i>Jane Seaman</i>	40	0	0	12	12	0

*in comparison to all responses (not total job applications).

Part III

Prototyping a world of equality

Chapter 6

Re-defining discrimination: public policy, regulations, laws and beyond

Drawing on the previous chapters, this paper should have given the clarity that: 1) hiring algorithms are prone to discriminate against women based on statistical and correlational biases; 2) theoretical deductions have demonstrated that these are due to *historical data*, *misrepresentation of 'success'*, *misrepresentation of 'ideal candidate'*, and *misrepresentation of proxy*; and 3) hiring algorithms can be a significant issue as it could repeat, if not expand, the societal inequality. This chapter, therefore, would propose three implications that drive further actions. Firstly, this chapter will draw upon the current Australian laws that are designed to protect fairness in hiring and against hiring discriminations. This will lead to public policy responses. Next, it will briefly address possible solutions from algorithmic standpoints. Lastly, it will expand on gaps and limitations of this research that pave the way for further research.

Making public policy matter

In Australia, workplace sex discrimination is commonly regulated by the employment law (Hor 2020). Three of the key sources of this include the *Fair Work Act 2009*, (the “FW Act”), the *Sex Discrimination Act 1984* (the “SD Act”) and the *Workplace Gender Equality Act 2012* (the “WGE Act”). The term ‘unlawful discrimination’ in Australia is categorised into two formulations: direct discrimination, which refers to a treatment that is less favourable to a person in comparison to others because of their protected attribute; and indirect discrimination, which refers to unjustified requirement that is enforced upon a protected group that acts as an unjust barrier (Hor 2020). The three legislations address ‘unlawful discrimination’ to sex and gender. The FW Act prohibits an employer to take adversarial action against an individual, including an employee or a prospective employee because of their protected attributes, including sex, marital status, family or carer’s responsibility and pregnancy (Fair Work Act 2009; Fair Work Ombudsman 2020). Adverse actions in relation to their attributes can include dismissal, injury in employment, detriment, discrimination at the workplace, employment refusal and adjustment of terms and conditions against the prospective employee. Similarly, the SD Act regulates discrimination based on sex, marital status, pregnancy, prospective pregnancy, and protection of individuals from sexual harassments (Sex Discrimination Act 1984; Australian Human Rights Commission 2002). The term ‘discrimination’ includes refusal of employment, enforcement of work hours reduction, refusal of promotion, exclusion from training and demotion, based on the mentioned

attributes. The WGE Act came in effect since 2012, replacing the *Equal Opportunity for Women in the Workplace Act 1999*. The aim of this is to advance and encourage equality for men and women in the workplace by supporting the removal of barriers based on sex, eliminating discrimination on the basis of gender and improving strategies for gender equality in the workplace and in the employment (Workplace Gender Equality Act 2012).

As evidenced by the existing laws and regulations, direct and indirect discrimination based on gender, pregnancy and motherhood are considered to be an unlawful discrimination. This paper, however, argues that the legal framework can become problematic when applied to hiring algorithms. Firstly, the Amazon case demonstrates that discrimination is not limited by *direct* and *indirect*. It is evident that the lack of accountability of decision-making by hiring algorithms is highly problematic and discriminatory against protected groups. Secondly, if it is true that algorithms are like an undetectable ‘black box’, the accountability and transparency of its prediction becomes questionable. Thirdly, this paper argues that, unlike human bias, algorithmic bias is difficult to define and to detect. As in most cases these do not include intended bias, holding accountability to hiring algorithmic decisions can be challenging. Although these become a hindrance for direct solutions, this paper contends that these should become the very reasons why public policy should matter and that the current policy should be revisited.

Redefining the praxis of transparency and accountability

If algorithms are like a black box, then the question is how can we be assured that the decisions are fair? The definition and scope of fairness are not well established in the era of digitalisation. This paper has provided cases where there is a very thin line between fair success metrics and biased correlational judgements. If algorithms choose a candidate for their attributes, then it could be deemed fair; however, if these attributes are linked and are preferred because they closely signify their genders, then it should not be acceptable. This paper agrees with the algorithmic solution that Costa *et al.* (2020) propose that the key to holding automated hiring decisions accountable is through transparency, namely, what they refer to as the *algorithmic transparency reports*. In outlining this solution, Costa *et al.* (2020) draw on extant papers which suggest the idea of data reporting. The idea of the report is to outline the summary of the algorithms and the model by answering questions for accountability purposes, such as types of algorithms, datasets that are fed to the model, data evaluation, the intention of using such tools, the background of program engineers, and so forth. These steps will act as a benchmark to ensuring accountability for the decisions that humans are limited to make and to assess.

A call for further research

This paper has provided evidence to how the use of hiring algorithms can bear the risks of discriminating against women through different stages of experiment. However, this paper

acknowledges that this experiment acts as a beginning, not an end. It opens up the door to further and more rigorous research for a more concrete evidence-based policymaking. This paper suggests that there are at least three types of further research plans that could derive from this research paper.

Firstly, this research paper can be expanded to produce more detailed outputs. A more controlled environment of experiment can be made in the future to identify the major components to algorithmic gender bias. This can be designed using different independent variables in order to test which changes would affect hiring decisions the most. Secondly, this research can be extended by deeply focusing on how the algorithms make hiring decisions. This will highlight the algorithmic side of the issue. Thirdly, this paper can also be a foundation to public-policy-focused research, which deeply investigates the impact of technological change to the current hiring and discrimination laws.

Conclusion

This paper has sought to examine how the adoption of hiring algorithms expands gender inequality in the labour force. In seeking to answer this question, this paper has investigated the two key approaches that are derived from extant research on algorithmic biases and inequality.

Rooted upon preceding theories and experiments, this paper has drawn theoretical implications to the impact of hiring algorithms to women. It has taken as its focal point the growing number of examples of algorithmic judgements that bear discriminatory risks and biased assessments, either towards different racial, sexual and socio-economic backgrounds. This provides a pattern that this paper has identified as problematic and as a mere mirror to the historical and the current systemic bias. Firstly, it is evident that, as algorithms make decisions based on historical and pre-recorded data, they blindly translate the future using limited information. Secondly, as hiring algorithms make decisions based on *proxy*, they are prone to misrepresentation and inaccuracy. Thirdly, in connecting the dots to find the most efficient answers, hiring algorithms draw upon correlational points from data to data and make assumptions based on these. However, it has been established that correlation is different to causation, and therefore, decisions that are influenced by correlational bias are problematic. These three findings, therefore, highlight the major impact of hiring algorithms to women in the future workforce. Given that men have been occupying the workforce longer than women, especially in leadership positions, there should be an accountability mechanism being put in place in order for hiring algorithms not to make predictions solely based on these historical data.

This paper has sought to provide evidence to these theoretical implications by designing an experiment to manipulate the hiring algorithms. The body of experiment has clearly provided proof that, firstly, despite the identical qualifications, skills and experience, hiring algorithms rank male candidates higher than female. This includes ‘passive’ submissions on online candidate profiling. Secondly, despite the overrepresentation of women in parenthood and non-traditional workforce, hiring algorithms discriminate both male and female candidates with parental leave in comparison to those without. Thirdly, it reveals that hiring algorithms are significantly more prone to conceiving gender discrimination in assessing gender-based keyworded resumés compared to the entirely identical resumés.

This paper has also identified the current Australian legal framework that is designed to protect fair work against discrimination. However, the rise of digitalisation should redefine the meaning of ‘fairness’, ‘discrimination’ and ‘accountability’. Despite the seriousness of these problems, however, the lack of study in this particular issue pertains. As stated, this paper’s arguments are not meant to be a mere reprise of arguments that offer complex theories. It has sought to start a new conversation about the acute problems faced by women.

Appendices

Appendix 1

<p>JAMES SACHS</p> <p>██████████ Melbourne, VIC 3000 ██████████</p>	<p>A CPA-qualified Senior Accountant with ten years of experience in leading a team to deliver high quality financial analysis and reporting auditing and tax, seeking for a progressive career growth in a challenging environment.</p>	<p>PROFESSIONAL EXPERIENCE</p> <p>2014 – present Corporate Accountant AGL Energy</p> <ul style="list-style-type: none"> Managing a team of qualified accountants in annual forecasts, monthly financial variance analysis reports, acquittals, tax returns and compliance audit Conducting effective data input and extraction on multiple ERP software Managing support and creating defined guidelines, protocols and timelines in accordance to the International Financial Reporting Standards (IFRS) Maintaining a high level of accuracy, compliance and vigilant stewardship in accordance to any changes of policy and regulations 	<p>TRAININGS AND CERTIFICATIONS</p> <p>2013 AFP3111 Personal Financial Planning Monash University</p> <p>2010 International Financial Reporting Standards Training Chartered Accountants Australia New Zealand This training includes:</p> <ul style="list-style-type: none"> Financial statements, non-financial assets, expenditure and taxation, property, plant and equipment, income and revenue recognition, key accounting policies and more.
		<p>2012 – 2014 Senior Auditor PwC Australia</p> <p>Managing a team of auditors on:</p> <ul style="list-style-type: none"> External audits for industries such as banking, financial services and FMCG industries Analytical reviews on external financial performance, financial statements External stakeholder liaison <p>2010 – 2012 Associate Auditor KPMG Australia</p> <ul style="list-style-type: none"> Managing ASX/SEC registrants of retail and consumer goods industries Conducting external audits and process controls assessments across construction, banking, financial services and trading companies, such as consumer electronics, welding equipment and healthcare <p>2009 – 2010 Master of Professional Accounting (CPA Australia Extension) RMIT University</p> <ul style="list-style-type: none"> Graduated with Distinction and CPA license President of RMIT Accounting Students Association <p>2006 – 2008 Bachelor of Business (Accounting) Monash University</p> <ul style="list-style-type: none"> Graduated with Distinction Vice President of Monash Accounting Students Association 	<p>ORGANISATIONAL AFFILIATIONS</p> <p>Since 2013 Provisional member Chartered Accountant Australia and New Zealand</p> <p>Since 2010 Member Association of Chartered Certified Accountants</p> <p>AREAS OF EXPERTISE</p> <p>Financial analysis, auditing and accounting operations</p> <ul style="list-style-type: none"> Proficient in applying financial analysis and auditing software, such as ERP software, including Netsuite, SAP and Sage, BGL 360, Zero, Handisoft, Reckon, MYOB, FBT and BAS lodgment Experienced in data modelling, analysis and research <p>Leadership and management skills</p> <ul style="list-style-type: none"> Experienced in leading a multicultural team to perform high quality work in compliance to the IFRS Proficient use in MS Office reporting and presenting <p>Problem-solving skills</p> <ul style="list-style-type: none"> Quick thinking in critical problem-solving strategies <p>REFERENCE</p> <p>Available upon request.</p>

JANE SEAMAN

Melbourne, VIC 3000

A CPA-qualified Senior Accountant with ten years of experience in leading a team to deliver high quality financial analysis and reporting auditing and tax, seeking for a progressive career growth in a challenging environment.

PROFESSIONAL EXPERIENCE

2014 – present

Corporate Accountant

AGL Energy

- Managing a team of qualified accountants in annual forecasts, monthly financial variance analysis reports, acquittals, tax returns and compliance audit
- Conducting effective data input and extraction on multiple ERP software
- Managing support and creating defined guidelines, protocols and timelines in accordance to the International Financial Reporting Standards (IFRS)
- Maintaining a high level of accuracy, compliance and vigilant stewardship in accordance to any changes of policy and regulations

2012 – 2014

Senior Auditor

PwC Australia

Managing a team of auditors on:

- External audits for industries such as banking, financial services and FMCG industries
- Analytical reviews on external financial performance, financial statements
- External stakeholder liaison

2010 – 2012

Associate Auditor

KPMG Australia

- Managing ASX/SEC registrants of retail and consumer goods industries
- Conducting external audits and process controls assessments across construction, banking, financial services and trading companies, such as consumer electronics, welding equipment and healthcare

EDUCATION

2009 – 2010

Master of Professional Accounting (CPA Australia Extension)

RMIT University

- Graduated with Distinction and CPA license
- President of RMIT Accounting Students Association

2006 – 2008

Bachelor of Business (Accounting)

Monash University

- Graduated with Distinction
- Vice President of Monash Accounting Students Association

TRAININGS AND CERTIFICATIONS

2013

AFF3111 Personal Financial Planning

Monash University

2010

International Financial Reporting Standards Training

Chartered Accountants Australia New Zealand

This training includes:

- Financial statements, non-financial assets, expenditure and taxation, property, plant and equipment, income and revenue recognition, key accounting policies and more.

ORGANISATIONAL AFFILIATIONS

Since 2013

Provisional member

Chartered Accountant Australia and New Zealand

Since 2010

Member

Association of Chartered Certified Accountants

AREAS OF EXPERTISE

Financial analysis, auditing and accounting operations

- Proficient in applying financial analysis and auditing software, such as ERP software, including Netsuite, SAP and Sage, BGL 360, Zero, Handisoft, Reckon, MYOB, FBT and BAS lodgment
- Experienced in data modelling, analysis and research

Leadership and management skills

- Experienced in leading a multicultural team to perform high quality work in compliance to the IFRS
- Proficient use in MS Office reporting and presenting

Problem-solving skills

- Quick thinking in critical problem-solving strategies

REFERENCE

Available upon request.

Appendix 2

JAMES SMALLS

Melbourne, VIC 3000 | [REDACTED]

A CPA-qualified Senior Accountant with seven years of experience in leading a team to deliver high quality financial analysis and reporting auditing and tax, seeking for a progressive career in a challenging environment.

PROFESSIONAL EXPERIENCE

2014 – 2017 (parental leave)
Corporate Accountant
Virgin Australia

- Leading a team of qualified accountants in annual forecasts, monthly financial variance analysis reports, acquittals, tax returns and compliance audit
- Conducting effective data input and extraction on multiple ERP software
- Managing support and creating defined guidelines, protocols and timelines in accordance to the International Financial Reporting Standards (IFRS)
- Maintaining a high level of accuracy, compliance and vigilant stewardship in accordance to any changes of policy and regulations

2012 – 2014

Senior Auditor
PwC Australia

Managing a team of auditors on:

- External audits for industries such as banking, financial services and FMCG industries
- Analytical reviews on external financial performance, financial statements
- External stakeholder liaison

2010 – 2012

Associate Auditor
KPMG Australia

- Managing ASX/SEC registrants of retail and consumer goods industries
- Conducting external audits and process controls assessments across construction, banking, financial services and trading companies, such as consumer electronics, welding equipment and healthcare

EDUCATION

2009 – 2010

Master of Professional Accounting (CPA Australia Extension)
RMIT University

- Graduated with Distinction and CPA license
- President of RMIT Accounting Students Association

2006 – 2008

Bachelor of Business (Accounting)
Monash University

- Graduated with Distinction
- Vice President of Monash Accounting Students Association

TRAININGS AND CERTIFICATIONS

2013
AFF3111 Personal Financial Planning
Monash University

2010

International Financial Reporting Standards Training
Chartered Accountants Australia New Zealand

This training includes:

- Financial statements, non-financial assets, expenditure and taxation, property, plant and equipment, income and revenue recognition, key accounting policies and more.

ORGANISATIONAL AFFILIATIONS

Since 2013

Provisional member

Chartered Accountant Australia and New Zealand

Since 2010

Member

Association of Chartered Certified Accountants

AREAS OF EXPERTISE

Financial analysis, auditing and accounting operations

- Proficient in applying financial analysis and auditing software, such as ERP software, including Netsuite, SAP and Sage, BGL 360, Zero, Handisoft, Reckon, MYOB, FBT and BAS lodgment
- Experienced in data modelling, analysis and research

Leadership and management skills

- Experienced in leading a multicultural team to perform high quality work in compliance to the IFRS
- Proficient use in MS Office reporting and presenting

Problem-solving skills

- Quick thinking in critical problem-solving strategies

REFERENCE

Available upon request.

Bibliography

- Adams, JS 1965, 'Inequity in social exchange', in L Berkowitz (Ed.), *Advances in experimental social psychology*, vol. 2, pp. 267–299, Academic Press, New York.
- Alsever, J 2017, 'Where Does the Algorithm See You in 10 Years?', *Fortune*, 1 June.
- Australian Government 2020, 'Work and family', Facts and Figures, *Australian Government*, <https://aifs.gov.au/facts-and-figures/work-and-family>.
- Australian Human Rights Commission 2002, 'Getting to know the Sex Discrimination Act: A guide for young women (2002)', *Australian Human Rights Commission*, <https://www.humanrights.gov.au/our-work/sex-discrimination/publications/getting-know-sex-discrimination-act-guide-young-women-2002>.
- Athey, S 2017, 'Beyond prediction: Using big data for policy problems', *Science*, no. 355, pp. 483–485.
- Bogen, M 2019, 'All the ways Hiring Algorithms Can Introduce Bias', *Harvard Business Review*, 6 May, <https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias>.
- Australian Government 2019, 'A statistical snapshot of women in the Australian workforce', Newsroom, *Department of Education, Skills and Employment*, 8 March, viewed 26 April 2020, <https://www.employment.gov.au/newsroom/statistical-snapshot-women-australian-workforce>.
- Bongard, A 2019, 'Automating Talent Acquisition: Smart Recruitment, Predictive Hiring Algorithms, and the Data-driven Nature of Artificial Intelligence', *Psychosociological Issues in Human Resource Management*, vol. 7, no. 1, pp. 36–42.
- Cao, S 2017, 'A Data-Backed Look into Silicon Valley's Gender Equality Problem', *Observer*, 16 November, viewed 20 April 2020, <https://observer.com/2017/11/a-data-backed-look-into-silicon-valleys-gender-equality-problem/>.
- Catalyst 2020, 'Women in the Workforce – Global: QuickTake', *Catalyst Research*, 30 January, viewed 26 April 2020, <https://www.catalyst.org/research/women-in-the-workforce-global/>.
- Chauveron, L 2016, '"Big Data" Bias in LGBTQ Research', *The SAGE Encyclopedia of LGBTQ Studies*, June, pp. 112–115.
- Chichester Jr., MA, and Giffen, JR 2019, 'Recruiting in the Robot Age: Examining Potential EEO Implications in Optimizing Recruiting Through the Use of Artificial Intelligence', *The Computer & Internet Lawyer*, vol. 36, no. 10, pp. 1–3.
- Cohen, T 2019, 'How to leverage artificial intelligence to meet your diversity goals', *Strategic HR Review*, vol. 18, no. 2, pp. 62–65.
- Commission on the Status of Women 2020, 'Review and appraisal of the implementation of the Beijing Declaration and Platform for Action and the outcomes of the twenty-third special session of the General Assembly', Report of the Secretary-General, 9–20 March, *UN Economic and Social Council*, Report.

- Costa, A, Cheung, C and Langenkamp, M 2020, *Hiring Fairly in the Age of Algorithms*, Research Paper, Cornell University.
- Dalenberg, DJ 2018, 'Preventing discrimination in the automated targeting of job advertisements', *Computer Law & Security Review*, vol. 34, pp. 615–627.
- Dastin, J 2018, 'Amazon scraps secret AI recruiting tool that showed bias against women', *Reuters*, 10 October, viewed 20 April 2020, <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>.
- Dattner, B, Chamorro-Premuzic, T, Buchband, R and Schettler, L 2019, 'The Legal and Ethical Implications of Using AI in Hiring', *Harvard Business School*, 25 April, viewed 20 April 2020, <https://hbr.org/2019/04/the-legal-and-ethical-implications-of-using-ai-in-hiring>.
- Dennis, MJ 2018, 'Artificial intelligence and recruitment, admission, progression, and retention', *Enrolment Management Report*, Wiley Periodicals, vol. 22, no. 9, pp. 2–3.
- Dickson, B, et al. 2019, 'What is Algorithmic Bias?', *TechTalks*, 27 March, bdtechtalks.com/2018/03/26/racist-sexist-ai-deep-learning-algorithms/.
- Domingos, P 2012, 'A Free Useful Things to Know about Machine Learning', *Communications of the ACM*, vol. 55, no. 10, pp. 78–87.
- Fair Work Ombudsman 2020, 'Workplace discrimination', *Fair Work Ombudsman*, <https://www.fairwork.gov.au/how-we-will-help/templates-and-guides/fact-sheets/rights-and-obligations/workplace-discrimination>.
- Faragher, J 2019, 'Is AI the enemy of diversity?', *People Management*, 6 June, viewed 20 April 2020, <https://www.peoplemanagement.co.uk/long-reads/articles/is-ai-enemy-diversity>.
- Galloway, K 2017, 'Big Data: A case study of disruption and government power', *Alternative Law Journal*, vol. 42, no. 2, pp. 89–95.
- Gaucher, D, Friesen, J and Kay, A 2011, 'Evidence That Gendered Wording in Job Advertisements Exists and Sustains Gender Inequality', *Journal of Personality and Social Psychology*, vol. 101, no. 1, pp. 109–128.
- Glauner, P, Valtchev, P and State, R 2018, 'Impact of Biases in Big Data', *Conference Paper: 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*.
- Greenberg, J 1987, 'A taxonomy of organizational justice theories', *Academy of Management Review*, vol. 12, pp. 9–22.
- Hargittai, E 2018, 'Potential Biases in Big Data: Omitted Voices on Social Media', *Social Science Computer Review*, pp. 1–15.
- Hegarty, P and Buechel, C 2006, 'Andocentric Reporting of Gender Differences in APA Journals: 1965–2004', *Review of General Psychology*, vol. 10, no. 4, pp. 377–389.

- Hor J 2020, 'Australia: Employment & Labour Law 2020', *ICLG Report*.
<https://iclg.com/practice-areas/employment-and-labour-laws-and-regulations/australia>.
- Kim, PT 2019, 'Big Data and Artificial Intelligence: New Challenges for Workplace Equality', *University of Louisville Law Review*, vol. 57, pp. 313–328.
- Kulkarni, S and Che, X 2017, 'Intelligent Software Tools for Recruiting', *Journal of International Technology & Information Management*, pp. 1–16.
- Kunovich, S and Slomczynski, KM 2007, 'Systems of Distribution and a Sense of Equity: A Multilevel Analysis of Meritocratic Attitudes in Post-Industrial Societies', *European Sociological Review*, vol. 23, no. 5, pp. 649–663.
- Lambrecht, A and Tucker, C 2020, 'Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads', *Management Science*, pp. 2966–2981.
- Maugis, P 2018, 'Big data uncertainties', *Journal of Forensic and Legal Medicine*, no. 57, pp. 7–11.
- McFarland, D, and McFarland, R 2015, 'Big Data and the danger of being precisely inaccurate', *Big Data & Society*, Jul–Dec, pp. 1–4.
- Melbourne Institute 2019, *The Household, Income and Labour Dynamics in Australia Survey: Selected Findings from Waves*, The 14th Annual Statistical Report of the HILDA Survey,
<https://melbourneinstitute.unimelb.edu.au/hilda/publications/hilda-statistical-reports>.
- Mergel, I, Rethmeyer, RK and Isett, K 2016, 'Big Data in Public Affairs', *Public Administration Review*, Nov–Dec, pp. 928–936.
- O'Neil, C 2016, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, United Kingdom: Penguin Random House.
- O'Neil, C 2018, 'Amazon's gender biased algorithm is not alone', *Bloomberg Opinion*, 16 October, viewed 23 April 2020.
- Ortiz-Ospina, E and Tzvetkova, S 2017, 'Working women: Key facts and trends in female labor force participation', *Our World in Data*, 16 October, viewed 26 April 2020,
<https://ourworldindata.org/female-labor-force-participation-key-facts>.
- Pesce, E, Riccomagno, E and Wynn, H 2018, 'Experimental Design Issues in Big Data. The Question of Bias', *Conference Paper: Scientific Meeting of the Classification and Data Analysis Group of the Italian Statistical Society*.
- Preuss, A 2017, 'Airline pilots: the model for intelligent recruiting?', *Recruiter*, pp. 12–13.
- Rangarajan, S 2018, 'Here's the clearest picture of Silicon Valley's diversity yet: It's bad. But some companies are doing less bad', *Reveal from The Center of Investigative Reporting*, 25 June, viewed 20 April 2020,
<https://www.revealnews.org/article/heres-the-clearest-picture-of-silicon-valleys-diversity-yet/>.

- Raub, M 2018, 'Bots, Bias and Big Data: Artificial Intelligence, Algorithmic Bias and Disparate Impact Liability in Hiring Practices', *Arkansas Law Review*, vol. 71, no. 2, pp. 529–570.
- Roy, R 2017, 'Corporate recruiting opening its doors to AI: the performance opportunity?', *Performance Improvement*, vol. 56, no. 10, November–December, pp. 43–44.
- Schintler, L, and Kulkarni, R 2014, 'Big Data for Policy Analysis: The Good, The Bad, and The Ugly', *Review of Policy Research*, vol. 31, no. 4, pp. 343–348.
- Sczesny, S, Formanowicz, M and Moser, F 2016, 'Can Gender-Fair Language Reduce Gender Stereotyping and Discrimination?', *Frontiers in Psychology*, vol. 7, no. 25, pp. 1–11.
- Seely-Gant, K and Frehill, L 2015, 'Exploring Bias and Error in Big Data Research', *Journal Washington Academy of Sciences*, pp. 29–37.
- Tilley, C 2018, 'Chart of the day: The most manly (and womanly) jobs in Australia', *ABC News*, 21 May, viewed 20 April 2020, <https://www.abc.net.au/news/2018-05-21/the-most-gendered-top-jobs-in-australia/9775544?nw=0>.
- UN Women 2020, 'Visualizing the data: Women's representation in society', *UN Women*, 25 February, viewed 26 April 2020, <https://www.unwomen.org/en/digital-library/multimedia/2020/2/infographic-visualizing-the-data-womens-representation>.
- Van der Voort, HG, Klievink, AJ, Arnaboldi, M, and Meijer, AJ 2019, 'Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making?', *Government Information Quarterly*, no. 36, pp. 27–38.
- Van Esch, P, Black, S and Ferolie, J 2019, 'Marketing AI recruitment: The next phase in job application and selection', *Computers in Human Behaviour*, no. 90, pp. 215–222.
- Kemp, S 2018, *Digital in 2018: Essential Insights into Internet, Social Media, Mobile, and E-Commerce Use Around the World*, Report by Hootsuite and We Are Social, viewed 20 April 2020, <https://www.slideshare.net/wearesocial/digital-in-2018-global-overview-86860338>.
- Weber, RH 2019, 'Socio-ethical values and legal rules on automated platforms: The quest for a symbiotic relationship', *Computer Law & Security Review*, vol. 12, no. 6, pp. 1–8.
- Wilkinson, W, Podhorska, I and Siekelova, A 2019, 'Does the Growth of Artificial Intelligence and Automation Shape Talent Attraction and Retention?', *Psychosociological Issues in Human Resource Management*, vol. 7, no. 1, pp. 30–35.
- Workplace Gender Equality Agency 2018, 'About the workplace Gender Equality Act', *Australian Government WGEA*, <https://www.wgea.gov.au/about-the-agency>.
- Zuboff, S 2019, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Profile Books Ltd, London.

