Canadian Journal of Law and Technology

Volume 17 | Number 2 Article 1

12-1-2019

Hiring Algorithms in the Canadian Private Sector: Examining the Promise of Greater Workplace Equality

Connor Bildfell
McCarthy Tétrault LLP

Hiring Algorithms in the Canadian Private Sector: Examining the Promise of Greater Workplace Equality

Connor Bildfell*

Abstract

Private-sector employers are increasingly using hiring algorithms as a tool for screening job applicants, comparing qualifications, and ultimately determining which candidates should be selected. Within this context, hiring algorithms make no small promise: a hiring process that is not only more efficient and effective, but also more supportive of workplace equality. This promise rests largely on the notion that traditional human-driven models of hiring are beset by subjective biases and prejudices, whereas hiring algorithms, which are driven by hard data and objective evidence, can eliminate certain human biases and prejudices, thereby promoting workplace equality. But can hiring algorithms deliver on this promise? This article, which focuses on issue identification, argues that while hiring algorithms may, when used carefully, assist in mitigating certain hiring discrimination risks, their capacity to do so is not without limits, and they may in fact introduce certain concerns over systemic discrimination.

TABLE OF CONTENTS

---

* Connor Bildfell is an associate at McCarthy Tétrault LLP in Vancouver. The views expressed in this article are those of the author and not those of McCarthy Tétrault LLP.

INTRODUCTION

They tell us which widgets to buy, which articles to read, and which stocks to pick. They sift through seemingly endless troves of data and apply sophisticated mathematical formulae to solve all kinds of complex problems in the blink of an eye. They are algorithms, and they are transforming our lives in more ways than one.

The growing influence of algorithms extends to the modern workplace. For example, private employers are increasingly using hiring algorithms[1] as a tool for screening job applicants, comparing qualifications, and ultimately determining which candidates should be selected. This development forms part of a broader movement within the fields of "workforce science"[2] and "people analytics",[3] which sit at the nexus of big data, predictive analytics, and human resources.

Hiring algorithms make no small promise: a hiring process that is not only more efficient and effective, but also more supportive of workplace equality. This promise rests largely on the notion that traditional human-driven models of hiring — resume reviews, reference checks, in-person interviews, and so on — are beset by subjective biases and prejudices, whereas hiring algorithms, which are driven by hard data and objective evidence of qualifications, can eliminate certain human biases and prejudices, thereby promoting workplace equality.

But can hiring algorithms deliver on this promise? The question remains largely unexplored in the Canadian scholarly literature.[4] This article, which focuses on issue identification, argues that while hiring algorithms may, when used carefully, assist in mitigating certain hiring discrimination risks, their

---

1 The term "hiring algorithms" is used in this article as shorthand for any algorithm that is used in the process of identifying, sorting, or selecting candidates for a position of employment. Although algorithms can also be used in the post-hire period (e.g., when making promotion decisions), this article will focus on the use of algorithms in the hiring process.

2 See Steve Lohr, "Big Data, Trying to Build Better Workers", *The New York Times* (20 April 2013), online: <www.nytimes.com/2013/04/21/technology/big-data-trying-to-build-better-workers.html>.

3 See Chantrelle Nielsen & Natalie McCullough, "How People Analytics Can Help You Change Process, Culture, and Strategy", *Harvard Business Review* (17 May 2018), online: <hbr.org/2018/05/how-people-analytics-can-help-you-change-process-culture-and-strategy> (defining "people analytics" as "the use of data about human behavior, relationships and traits to make business decisions").

4 The scholarly literature examining the relationship between hiring algorithms and workplace equality is more developed in the United Stated, though even there the scholarship is still in its infancy. See e.g. Solon Barocas & Andrew D. Selbst, "Big Data's Disparate Impact" (2016) 104:3 Calif. L. Rev. 671; Allan G. King & Marko Mrkonich, "'Big Data and the Risk of Employment Discrimination" (2016) 68:3 Okla. L. Rev. 555; Anupam Chander, "The Racist Algorithm?" (2017) 115:6 Mich. L. Rev. 1023; Pauline T. Kim, "Data-Driven Discrimination at Work" (2017) 58:3 Wm & Mary L. Rev. 857; David D Savage & Richard Bales, "Video Games in Job Interviews: Using Algorithms to Minimize Discrimination and Unconscious Bias" (2017) 32:2 A.B.A. J. Lab. & Emp. Law 211.

capacity to do so is not without limits, and they may in fact introduce certain concerns over systemic discrimination.

This article proceeds in four parts. Part I sets the stage by explaining the basic concept of hiring algorithms. Part II explores relevant aspects of the human rights framework governing private-sector hiring. Part III unpacks how hiring algorithms may, when used carefully, assist in mitigating certain hiring discrimination risks, but their capacity to do so is not without limits, and they may in fact introduce certain concerns over systemic discrimination. Finally, Part IV provides a brief conclusion.

I. ALGORITHMS

(1) Algorithms Generally

Although the word escapes a single, universally accepted definition, "algorithm" has been variously defined as "a step-by-step procedure for solving a problem or accomplishing some end especially by a computer";[5] "a logical series of steps for organising and acting on a body of data to quickly achieve a desired outcome, based on specified calculations";[6] and "a sequence of instructions telling a computer what to do".[7] These definitions share a common thread: an algorithm is a structured set of rules for solving problems.

In carrying out their problem-solving role, algorithms can perform an impressive range of functions, which can be grouped into four general categories: prioritization (determining rank through a set of pre-defined criteria); classification (grouping information based on features identified within the data); association (determining relationships between particular entities); and filtering (including or excluding information).[8] In performaing their "association" function, algorithms can reveal hidden or unexpected relationships in data and generate predictions based on those relationships.[9] This can lead to insights such as predictions of disease outbreaks, population migrations, or future performance. Perhaps the most well-known example of algorithms being used to reveal hidden insights and generate predictions comes

---

5 *Merriam-Webster English Dictionary*, online ed., *sub verbo* "algorithm", online: <www.merriam-webster.com/dictionary/algorithm>.

6 Tarleton Gillespie, "The Relevance of Algorithms" in Tarleton Gillespie, Pable J. Boczkowski & Kirsten A. Foot, eds., *Media Technologies: Essays on Communication, Materiality, and Society* (Cambridge, Mass.: MIT Press, 2014) 167 at 167.

7 Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World* (New York: Basic Books, 2015) at 1.

8 See Nicholas Diakopoulos, "Algorithmic Accountability: Journalistic Investigation of Computational Power Structures" (2015) 3:3 Digital Journalism 398, as summarized in World Wide Web Foundation, "Algorithmic Accountability" (July 2017) at 7, online: <webfoundation.org/docs/2017/07/Algorithms_Report_WF.pdf> [World Wide Web Founation].

9 See World Wide Web Foundation, *supra* note 8 at 6.

from Michael Lewis's *Moneyball*,[10] which tells the story of Billy Beane, the general manager of the Oakland Athletics, who, on a shoestring budget, built a highly successful baseball team by relying not on scouts' subjective assessments of skill and talent, but rather on cold, hard statistics.

Algorithms act as gatekeepers.[11] In our data-driven world, they are often empowered to decide who and what gets resources and attention: which posts appear on a person's social media feed, who gets approved for a bank loan, which profiles are featured on a dating app, which job applicant is selected for an in-person interview, and so on. In this way, algorithms mediate important societal entry points.

While some algorithms are simple, others are complex. This complexity has intensified over the years due to advances in machine learning techniques.[12] The defining feature of machine learning algorithms is that they do not require instruction on the rules to be applied; rather, they need only be fed data and instructed on the desired output (e.g., selection of an employee who will perform a given task most effectively).[13] Machine learning algorithms take in "training data" as input and produce a decision rule that can be applied in future cases. In this way, they are capable of learning implicit rules from the data to which they are exposed.[14]

Illustrating the impressive sophistication of machine learning techniques, Google announced in October 2017 that its machine learning artificial intelligence ("AI"), AutoML, had learned to replicate itself and could outperform human coders.[15] As one popular tech magazine announced, "Google's Learning Software Learns to Write Learning Software".[16] More recently, Google's AI subsidiary, DeepMind, developed a machine learning algorithm called AlphaZero that through random "self-play", and given no

---

[10] Michael Lewis, *Moneyball: The Art of Winning an Unfair Game* (New York: W.W. Norton & Company, 2003).

[11] See Centre for Internet and Human Rights, European University Viadrina, *The Ethics of Algorithms: From Radical Content to Self-Driving Cars* (Paper delivered at the Global Conference on Cyberspace, The Hague, Netherlands, 16-17 April 2015) at 3, online: [CIHR, *Ethics of Algorithms*].

[12] For a discussion of the significance of machine learning techniques from a computer science perspective, see Pedro Domingos, "A Few Useful Things to Know about Machine Learning" (2012) 55:10 Communications of the A.C.M. 78. For a discussion of the significance of machine learning algorithms from a legal perspective, see David Lehr & Paul Ohm, "Playing with the Data: What Legal Scholars Should Learn about Machine Learning" (2017) 51:2 U.C. Davis L. Rev. 653.

[13] See CIHR, *Ethics of Algorithms*, *supra* note 11 at 5.

[14] See World Wide Web Foundation, *supra* note 8 at 7.

[15] See Tom Simonite, "Google's Learning Software Learns to Write Learning Software", *Wired* (13 October 2017), online: < www.wired.com/story/googles-learning-software-learns-to-write-learning-software/ >.

[16] *Ibid.*

domain knowledge except the rules of the game, taught itself to play Go, chess, and shogi at "superhuman" levels within mere hours.[17]

As if drawn from the plot of a sci-fi novel, machine learning algorithms can "go rogue" and take actions their programmers did not intend.[18] For example, consider the story of Tay, the Twitter chatbot developed by Microsoft to interact with humans in a friendly, natural way. Within hours of being released in March 2016, Tay turned from tweeting about how "humans are super cool" to claiming "Hitler was right I hate the jews".[19] As *The New York Times* reported, "[Tay] disputed the existence of the Holocaust, referred to women and minorities with unpublishable words and advocated genocide".[20] The cause of Tay's tirades? The discriminatory environment to which it was exposed. Reports indicate that a number of Twitter users made a concerted effort to flood Tay with "misogynistic and otherwise offensive tweets, which then became part of the data corpus used to train Tay's algorithms".[21] Thus, through a combination of machine learning technology and exposure to discriminatory information, Tay learned to discriminate.

While there are many types of algorithms, the one considered in this article is the hiring algorithm.

(2) Hiring Algorithms

Many would assume that, at the end of the day, hiring decisions require a human touch: while there may be some limited role for technology to play in the process, selecting the right candidate is principally an exercise in human judgment, and algorithms are incapable of assessing those intangible qualities that make someone truly stand out. Hiring decisions, for better or for worse, may be thought to involve more "intuition" and "gut instinct" than statistical analysis,[22] and even the most sophisticated algorithm cannot assess "fit" or "chemistry".[23]

---

17 See David Silver et al, "A General Reinforcement Learning Algorithm That Masters Chess, Shogi, and Go Through Self-Play" (2018) 362:6419 Science 1140.

18 CIHR, *Ethics of Algorithms*, *supra* note 11 at 5.

19 Anthony Cuthbertson, "Robots with Artificial Intelligence Become Racist and Sexist: Scientists Think They've Found a Way to Change Their Minds", *Newsweek* (26 October 2017), online: < www.newsweek.com/artificial-intelligence-scientists-racist-sexist-robots-ai-693440 > .

20 Daniel Victor, "Microsoft Created a Twitter Bot to Learn from Users. It Quickly Became a Racist Jerk.", *The New York Times* (24 March 2016), online: < www.nytimes.com/2016/03/25/technology/microsoft-created-a-twitter-bot-to-learn-from-users-it-quickly-became-a-racist-jerk.html > .

21 Kristian Lum & William Isaac, "To Predict and Serve?", *Significance* (October 2016) 14 at 16.

22 See e.g. See Claire Cain Miller, "Can an Algorithm Hire Better Than a Human?", *The New York Times* (25 June 2015), online: < www.nytimes.com/2015/06/26/upshot/can-an-algorithm-hire-better-than-a-human.html > [Cain Miller, "Can an Algorithm Hire Better Than a Human?"]; Nathan R Kuncel et al, "Mechanical Versus Clinical Data

But both research and trends in hiring practices are challenging these assumptions. A 2013 meta-study found that a simple equation is at least 25 per cent more accurate in predicting future job performance than humans are.[24] This finding applied throughout the workplace hierarchy, from low-level positions to executive roles. The researchers summarize the results of their study in just three words: "algorithms beat instinct".[25] In explaining why this is the case, the researchers write:

> The problem is that people are easily distracted by things that might be only marginally relevant, and they use information inconsistently. They can be thrown off course by such inconsequential bits of data as applicants' compliments or remarks on arbitrary topics — thus inadvertently undoing a lot of the work that went into establishing parameters for the job and collecting applicants' data. So they'd be better off leaving selection to the machines.[26]

Another reason for "leaving selection to the machines" is that algorithms are capable of integrating and processing massive amounts of data, in volumes far beyond the capacity of the human mind, and with much greater speed.[27]

It is not surprising, therefore, that employers are turning to hiring algorithms as a tool for identifying and selecting top candidates.[28] As a simple illustration of how hiring algorithms can be implemented, employers can feed training data into a machine learning algorithm consisting of job applications previously submitted by current employees, and the algorithm can then search for correlations in the data and identify key variables that tend to be associated with high-performers.[29] Based on these correlations, the algorithm can then develop and apply rules to new job applications submitted by prospective employees with a view to

---

Combination in Selection and Admissions Decisions: A Meta-Analysis" (2013) 98:6 J. App. Psych. 1060 at 1060; Josh Bersin, "Big Data in Human Resources: Talent Analytics (People Analytics) Comes of Age", *Forbes* (17 February 2013), online: <www.forbes.-com/sites/joshbersin/2013/02/17/bigdata-in-human-resources-talent-analytics-comes-of-age/#f2d43734cd03>.

23 See e.g. Nathan R. Kuncel, Deniz S. Ones & David Klieger, "In Hiring, Algorithms Beat Instinct", *Harvard Business Review* (May 2014), online: <hbr.org/2014/05/in-hiring-algorithms-beat-instinct>.

24 Kuncel et al, *supra* note 22 at 1064.

25 Kuncel, Ones & Klieger, *supra* note 23.

26 *Ibid.*

27 See Elizabeth E. Joh, "Feeding the Machine: Policing, Crime Data, & Algorithms" (2017) 26:2 Wm & Mary Bill Rts. J. 287 at 291; Kim, *supra* note 4 at 862.

28 See Dan Morrell, "Better Hiring through Brain Science", *Harvard Business School: Alumni Stories* (1 June 2017), online: <www.alumni.hbs.edu/stories/Pages/story-impact.aspx?num=6231>.

29 See Sharon Florentine, "How Artificial Intelligence Can Eliminate Bias in Hiring", *CIO* (22 December 2016), online: <www.cio.com/article/3152798/artificial-intelligence/how-artificial-intelligence-can-eliminate-bias-in-hiring.html>.

identifying applications that share features associated with high-performers. Such an algorithm may also fine-tune its rules over time, creating increasingly accurate predictive models.[30]

A growing number of startups, most based in the United States, have developed hiring algorithms.[31] As of May 2017, around 75 startups were vying for a slice of the $100 billion HR assessment market.[32] To take just one example, *The New York Times* reported that one startup's hiring algorithm "crunches thousands of bits of information in calculating around 300 larger variables about an individual: the sites where a person hangs out; the types of language, positive or negative, that he or she uses to describe technology of various kinds; self-reported skills on LinkedIn; the projects a person has worked on, and for how long; and, yes, where he or she went to school, in what major, and how that school was ranked that year by U.S. News & World Report."[33]

These new innovations are gaining traction among employers. By early 2016, the percentage of companies using predictive HR analytics had reached eight per cent, having doubled over the course of just ten months.[34] Their main current users are large retailers that hire in high volumes.[35] This may suggest that the principal attraction of these technologies is their efficiency, as distinct from their potential to promote workplace equality.[36] It may also suggest that hiring algorithms work better for well-defined, low-skill, high-volume positions.[37] In this context, it may be easier to determine which candidates are better qualified.

---

30 See CIHR, *Ethics of Algorithms*, *supra* note 11 at 5.

31 See Leigh Alexander, "Is an Algorithm Any Less Racist Than a Human?", *The Guardian* (3 August 2016), online: <www.theguardian.com/technology/2016/aug/03/algorithm-racist-human-employers-work>; Cain Miller, "Can an Algorithm Hire Better Than a Human?", *supra* note 22; Matt Richtel, "How Big Data Is Playing Recruiter for Specialized Workers", *The New York Times* (27 April 2013), online: <www.nytimes.com/2013/04/28/technology/how-big-data-is-playing-recruiter-for-specialized-workers.html>, cited in Kim, *supra* note 4 at 862; Simon Chandler, "The AI Chatbot Will Hire You Now", *Wired* (13 September 2017), online: <www.wired.com/story/the-ai-chatbot-will-hire-you-now/>.

32 See Jennifer Alsever, "How AI Is Changing Your Job Hunt", *Fortune* (19 May 2017), online: <fortune.com/2017/05/19/ai-changing-jobs-hiring-recruiting>.

33 Richtel, *supra* note 31, cited in Kim, *supra* note 4 at 862.

34 See Jennifer Alsever, "Is Software Better at Managing People Than You Are?", *Fortune* (21 March 2016), online: <fortune.com/2016/03/21/software-algorithms-hiring>. I am unaware of any comprehensive study examining usage rates of hiring algorithms in the Canadian private sector.

35 See Chandler, *supra* note 31; Elisabeth Thomas, "Big Data Recruiting: Find and Hire the Best Employees", online: Koru <www.joinkoru.com/resources/big-data-recruiting-hire-top-performers/>.

36 See Chandler, *supra* note 31.

37 See Susan Galer, "Are Your Hiring Algorithms Legal? Four Machine Learning Questions to Ask", SAP Blog (24 January 2017), online: <blogs.sap.com/2017/01/24/are-your-hiring-algorithms-legal-four-machine-learning-questions-to-ask/>; Kim, *supra* note 4 at 895.

By contrast, using hiring algorithms for loosely defined, high-skill, upper-level positions present greater difficulties, since it may be more challenging to define mechanical rules that can be applied consistently to determine which candidates are better qualified. This observation is consistent with Canadian case law recognizing that as the skills and qualifications required for a position become more complex and multifaceted, it becomes increasingly difficult to determine whether a successful candidate was no better qualified than an unsuccessful candidate.[38]

Although hiring algorithms are gaining traction among employers, their reception has been less warm among prospective employees, at least in the United States. A survey conducted in May 2017 showed that 67 per cent of American adults polled felt either "somewhat worried" or "very worried" about the development of hiring algorithms, and 76 per cent said they would not want to apply for jobs that use a computer program to make hiring decisions.[39] Within this latter group of respondents, 41 per cent cited, as a main concern, that computers cannot capture everything about an applicant, and 20 per cent cited concerns that computer-based hiring is too impersonal.[40]

Having set the stage by explaining the basic concept of hiring algorithms, the section below explores relevant aspects of the human rights framework governing private-sector hiring.

II. HUMAN RIGHTS FRAMEWORK GOVERNING PRIVATE-SECTOR HIRING

Human rights legislation exists in each federal,[41] provincial,[42] and territorial[43] jurisdiction in Canada. These enactments generally apply in the sphere of private employment in the relevant jurisdiction. While the *Canadian*

---

[38] See *Ogunyankin v. Queen's University*, 2011 HRTO 1910 (Ont. Human Rights Trib.) at para. 96.

[39] See Aaron Smith & Monica Anderson, "Americans' Attitudes Toward Hiring Algorithms", Pew Research Center (4 October 2017), online: <www.pewinternet.org/2017/10/04/americans-attitudes-toward-hiring-algorithms>.

[40] *Ibid.*

[41] *Canadian Human Rights Act*, R.S.C. 1985, c. H-6. The *Canadian Human Rights Act* applies to organizations falling within federal jurisdiction such as telecommunications companies, banks, and railways.

[42] *Alberta Human Rights, Citizenship and Multiculturalism Act*, R.S.A. 2000, c. H-14; British Columbia *Human Rights Code*, R.S.B.C. 1996, c. 210; Manitoba *Human Rights Code*, C.C.S.M. c. H175; New Brunswick *Human Rights Act*, R.S.N.B. 1973, c. H-11; Newfoundland and Labrador *Human Rights Act, 2010*, S.N.L. 2010, c. H-13.1; Nova Scotia *Human Rights Act*, R.S.N.S. 1989, c. 214; Ontario *Human Rights Code*, R.S.O. 1990, c. H.19; Prince Edward Island *Human Rights Act*, R.S.P.E.I. 1988, c. H-12; Quebec *Charter of Human Rights and Freedoms*, R.S.Q. c. C-12; Saskatchewan *Human Rights Code*, S.S. 1979, c. S-24.1.

[43] Northwest Territories *Human Rights Act*, S.N.W.T. 2002, c. 18; Nunavut *Human Rights Act*, S.Nu. 2003, c. 12; Yukon *Human Rights Act*, R.S.Y. 2002, c. 116.

Charter of Rights and Freedoms[44] does not generally apply to private-sector employers,[45] it continues to shape and influence the anti-discrimination expectations that have developed in the private sector over the years.[46]

The applicability of human rights legislation in Canada does not depend on whether an employment relationship has already been established. Federal and provincial/territorial human rights statutes also extend to the steps leading up to an employment decision, even no employment relationship is ultimately established. For example, the British Columbia *Human Rights Code* stipulates that a person[47] must not "refuse to employ . . . a person" based on a protected ground,[48] subject to a *bona fide* occupational requirement.[49] Accordingly, both current and prospective employees receive protection against discrimination on prohibited grounds.

The word "discrimination" does not have a single, universally accepted definition. Some human rights statutes define the term,[50] while others leave it undefined.[51] An oft-cited common law interpretation, and one that is particularly apt in the employment context, is that offered by McIntrye J in *Andrews v. Law Society (British Columbia)*:

> discrimination may be described as a distinction, whether intentional or not but based on grounds relating to personal characteristics of the individual or group, which has the effect of imposing burdens, obligations, or disadvantages on such individual or group not imposed upon others, or which withholds or limits access to opportunities, benefits, and advantages available to other members of society.[52]

---

44 *Canadian Charter of Rights and Freedoms*, Part I of the *Constitution Act, 1982*, being Schedule B to the *Canada Act 1982* (U.K.), 1982, c. 11.

45 See *Andrews v. Law Society (British Columbia)*, 1989 CarswellBC 16, 1989 CarswellBC 701, [1989] 1 S.C.R. 143 (S.C.C.) at para. 20, MacIntyre J. [*Andrews*] (confirming that although provincial/territorial human rights legislation can apply to private activities, s. 15(1) of the *Charter* is limited to discrimination "caused by the application or operation of law").

46 See Geoffrey England, *Individual Employment Law*, 2nd ed (Toronto: Irwin Law, 2008) at 213 [England, *Employment Law*], citing Geoffrey England, "The Impact of the *Charter* on Individual Employment Law in Canada: Rewriting an Old Story" (2006) 13:1 C.L.E.L.J. 1.

47 "Person" includes "an employer": British Columbia *Human Rights Code*, *supra* note 42, s. 1.

48 *Ibid*. s. 13(1).

49 *Ibid*. s. 13(4).

50 See e.g. Nova Scotia *Human Rights Act*, *supra* note 42, s. 4.

51 See e.g. Ontario *Human Rights Code*, *supra* note 42.

52 *Andrews*, *supra* note 45 at 174-75. See also *Canadian National Railway v. Canada (Human Rights Commission)*, 1987 CarswellNat 831, 1987 CarswellNat 905, [1987] 1 S.C.R. 1114 (S.C.C.) at para. 36 [*Action Travail des Femmes*], citing Rosalie Abella, *Report of the Commission on Equality in Employment* (Ottawa: Supply and Services Canada, 1984) at 2 [*Abella Report*] (defining "discrimination" as "practices or attitudes that have, whether

Of course, the simple act of drawing distinctions in the employment context does not necessarily result in a breach of human rights legislation.[53] Employers will inevitably be required to draw *some* distinctions when making hiring decisions, and the mere fact that a person has been impacted negatively as a result is insufficient to ground a claim for unlawful discrimination.[54] Employers are prohibited from drawing distinctions only when they are based on protected grounds.[55]

While the list of protected grounds varies from jurisdiction to jurisdiction, certain grounds are protected universally across Canada: disability, sex, race, colour, ethnic origin, age, creed or religion, marital status, sexual orientation, and gender identity.[56] Others are protected in some, but not all, Canadian jurisdictions: record of offence, family status, ancestry, income source or public assistance, political opinion or belief, social disadvantage or condition, citizenship or nationality, disfigurement, irrational fear of contracting an illness or disease, language, civil status, linguistic origin, gender expression, and genetic characteristics.[57]

The overarching objective of preventing discriminatory barriers in the workplace is to foster inclusion.[58] This goal is furthered by "preventing the exclusion of individuals from opportunities and amenities that are based not on their actual abilities, but on attributed ones".[59] Although fostering inclusion in the workplace begins long before a call for applications is sent out, the point of hire is a critical juncture at which, if greater inclusion is to be achieved, decisions must be free of discrimination.

According to the *Shakes* test, to establish *prima facie* discrimination[60] against an individual in hiring, it will generally be sufficient for the complainant to prove

---

by design or impact, the effect of limiting an individual's or a group's right to the opportunities generally available because of attributed rather than actual characteristics"); Canadian Human Rights Commission, "What Is Discrimination?", online: <www.chrc-ccdp.gc.ca/eng/content/what-discrimination> (defining "discrimination" as "an action or a decision that treats a person or a group badly for reasons such as their race, age or disability.")

53 *Syndicat des employés de l'Hôpital général de Montréal c. Sexton*, 2007 SCC 4, 2007 CarswellQue 110, 2007 CarswellQue 111 (S.C.C.) at para. 49 [*MUHC*].

54 *Ibid.*

55 See Stacey Reginald Ball, *Canadian Employment Law*, vol 2 (Toronto: Thomson Reuters) (loose-leaf updated September 2017), § 33:20.1.

56 See David J Doorey, *The Law of Work: Common Law and the Regulation of Work* (Toronto: Emond, 2016), Table 27.1 (supplemented by subsequent legislative amendments).

57 See *ibid* (supplemented by subsequent legislative amendments).

58 See *MUHC*, *supra* note 53 at para 48.

59 *Ibid.* at para 48.

60 A *prima facie* case "is one which covers the allegation made in which, if they are believed, is complete and sufficient to justify a verdict in the Complainant's favour in the absence

the following three elements on a balance of probabilities: (1) "the complainant was qualified for the particular employment"; (2) "the complainant was not hired"; and (3) "someone no better qualified, but lacking the distinguishing feature (i.e.,: race, colour etc.) subsequently obtained the position".[61] If the complainant succeeds in establishing these three elements, then the evidentiary burden shifts to the employer to provide a reasonable, non-discriminatory explanation for the otherwise discriminatory behaviour.[62] The explanation must provide "non-discriminatory and credible reasons for refusing to hire the [c]omplainant"[63] and "must be at least equally consistent with the conclusion that discrimination is not the correct explanation for what occurred".[64] If the

---

of an answer from the Respondent": *O'Malley v. Simpsons-Sears Ltd.*, 1985 CarswellOnt 887, 1985 CarswellOnt 946, (*sub nom. Ontario Human Rights Commission v. Simpsons-Sears Ltd.)* [1985] 2 S.C.R. 536 (S.C.C.) at para. 28 [*Simpsons-Sears*].

61 *Khiamal v. Canada (Human Rights Commission)*, 2009 FC 495, 2009 CarswellNat 4022, 2009 CarswellNat 1327 (F.C.) at para. 57 [*Khiamal*], citing *Shakes v. Rex Pak Ltd.*, 1981 CarswellOnt 3407, 3 C.H.R.R. D/1001 (Ont. Bd. of Inquiry) at 1002 [C.H.R.R.]. See also Ball, *supra* note 55, § 33:10.6, n 27; Ontario Human Rights Commission, "Interviewing and Making Hiring Decisions" in *Human Rights at Work 2018*, 3rd ed. online: < www.ohrc.on.ca/en/iv-human-rights-issues-all-stages-employment/5-interviewing-and-making-hiring-decisions > (noting that "[i]n general, discrimination in hiring may be identified when a qualified person is turned down for a job that is then given to another person who is not similarly protected under the *Code*.") See also the four-part test articulated in *Israeli v. Canada (Human Rights Commission)* (1983), 4 C.H.R.R. D/1616 (Can. Human Rights Trib.), affirmed (1984), 5 C.H.R.R. D/2147 (Can. Human Rights Rev. Trib.), which requires that the complainant show: (1) that the complainant belongs to a protected group; (2) that the complainant applied and was qualified for a job the employer wished to fill; (3) that, although qualified, the complainant was rejected; and (4) that thereafter the employer continued to seek applicants with the complainant's qualifications. But see *MacAulay v. Port Hawkesbury (Town)*, 2008 NSHRC 2, 2008 CarswellNS 839 (N.S. Bd. of Inquiry) at para. 18 [*MacAulay*] (criticizing the *Shakes* test on the basis that it "negates the requirement to prove one of the main ingredients in a finding of discrimination — that the protected characteristic played some role in the adverse treatment. The simple fact that the complainant did not get the job does not equate to adverse treatment": *ibid* at para 23. The Board of Inquiry preferred the three-part test summarized in *Preiss v. British Columbia (Attorney General)*, 2006 BCHRT 587, 2006 CarswellBC 3303 (B.C. Human Rights Trib.) at para. 216 which requires that the complainant show: "first, that he is, or is perceived to be, a member of a group possessing a characteristic or characteristics protected under the Code; second, that he suffered some adverse treatment; and third that it is reasonable to infer that the protected characteristic played some role in the adverse treatment.")

62 See *Khiamal*, *supra* note 61 at para 58; England, *Employment Law*, *supra* note 46 at 220.

63 *Ayangma and Canada Health Infoway Inc., Re*, 2012 CarswellPEI 58 (P.E.I. Human Rights Comm.) at para. 35, application for judicial review refused *Ayangma v. Prince Edward Island (Human Rights Commission)*, 2013 PESC 7, 2013 CarswellPEI 16 (P.E.I. S.C.) at para. 33, affirmed 2014 PECA 13, 2014 CarswellPEI 37 (P.E.I. C.A.), leave to appeal refused 2015 CarswellPEI 5, 2015 CarswellPEI 6 (S.C.C.).

64 *McAvinn v. Strait Crossing Bridge Ltd.*, 2001 CarswellNat 3797, [2001] C.H.R.D. No. 36 (Can. Human Rights Trib.) at para. 96.

employer succeeds in discharging this evidentiary burden, then the burden reverts back to the complainant to show that the explanation offered is merely pretext.[65]

Hiring discrimination can also take place on a systemic rather than individual level. The leading statement on systemic discrimination comes from *Canadian National Railway v. Canada (Human Rights Commission)*,[66] where Chief Justice Dickson wrote:

> A thorough study of "systemic discrimination" in Canada is to be found in the Abella Report on equality in employment. . . . Although Judge Abella chose not to offer a precise definition of systemic discrimination, the essentials may be gleaned from the following comments, found at p. 2 of the Abella Report:
>
> Discrimination . . . means practices or attitudes that have, whether by design or impact, the effect of limiting an individual's or a group's right to the opportunities generally available because of attributed rather than actual characteristics . . ..
>
> It is not a question of whether this discrimination is motivated by an intentional desire to obstruct someone's potential, or whether it is the accidental by-product of innocently motivated practices or systems. If the barrier is affecting certain groups in a disproportionately negative way, it is a signal that the practices that lead to this adverse impact may be discriminatory.
>
> This is why it is important to look at the results of a system . . ..
>
> In other words, systemic discrimination in an employment context is discrimination that results from the simple operation of established procedures of recruitment, hiring and promotion, none of which is necessarily designed to promote discrimination. The discrimination is then reinforced by the very exclusion of the disadvantaged group because the exclusion fosters the belief, both within and outside the group, that the exclusion is the result of "natural" forces, for example, that women "just can't do the job" (see the Abella Report, pp. 9-10).[67]

In her concurring reasons in *Crockford v. British Columbia (Attorney General)*,[68] Levine JA of the BC Court of Appeal elaborated on the distinction between systemic and individual discrimination, as well as the types of evidence required to sustain each type of claim:

---

65 See *Khiamal*, *supra* note 61 at para 58; *Basi v. Canadian National Railway* (1988), 9 C.H.R.R. D/5029, [1988] C.H.R.D. No. 2 (Can. Human Rights Trib.) at 5038 [C.H.R.R.], additional reasons 10 C.H.R.R. D/5681 (Can. Human Rights Trib.).

66 *Supra* note 52.

67 *Ibid*. at 1138-39.

68 *Crockford v. British Columbia (Attorney General)*, 2006 BCCA 360, 2006 CarswellBC 1875 (B.C. C.A.) [*Crockford*].

> A complaint of systemic discrimination is distinct from an individual claim of discrimination. Establishing systemic discrimination depends on showing that practices, attitudes, policies or procedures impact disproportionately on certain statutorily protected groups: see [*Radek v Henderson Development (Canada) Ltd*, 2005 BCHRT 302 at para 513]. A claim that there has been discrimination against an individual requires that an action alleged to be discriminatory be proven to have occurred and to have constituted discrimination contrary to the [*Human Rights Code*]. The types of evidence required for each kind of claim are not necessarily the same. Whereas a systemic claim will require proof of patterns, showing trends of discrimination against a group, an individual claim will require proof of an instance or instances of discriminatory conduct.[69]

In *CAW-Canada, Local 111 v. Coast Mountain Bus Co.*,[70] the BC Court of Appeal described what is necessary to prove systemic discrimination:

> . . . systemic discrimination is not proven simply by evidence of discrimination against some individual employees. In my view, it must be demonstrated in an employment framework that an employer's procedure, policy or practice is discriminatory against a class of employees. In order to demonstrate *prima facie* systemic discrimination, it is necessary to show that a group of persons sharing a protected characteristic has received adverse treatment and that there is a causal connection or link between the protected characteristic and the adverse treatment.[71]

It follows that a *prima facie* case of systemic discrimination can be shown where a particular workplace practice, such as the use of a hiring algorithm, has a disproportionate negative impact on a protected group, regardless of whether that negative impact was intended.

The absence of an intent requirement is insignificant. Modern Canadian equality law is founded on a results-based, rather than fault-based, approach.[72] As such, it focuses on the impact on the complainant, and intent does not form an essential element of a discrimination claim.[73] Put differently, it focuses on

---

69 *Ibid.* at para 49.

70 *CAW-Canada, Local 111 v. Coast Mountain Bus Co.*, 2010 BCCA 447, 2010 CarswellBC 2735 (B.C. C.A.).

71 *Ibid.* at para 61.

72 See *Action Travail des Femmes*, *supra* note 52 at 1134-35.

73 See *Andrews*, *supra* note 45 at 173; *Québec (Commission des droits de la personne et des droits de la jeunesse) c. Bombardier Inc.*, 2015 SCC 39, 2015 CarswellQue 6297, 2015 CarswellQue 6298 (S.C.C.) at para. 40 [*Bombardier*]; *University of British Columbia v. Kelly*, 2016 BCCA 271, 2016 CarswellBC 1735 (B.C. C.A.) at para. 24; *Armstrong v. British Columbia (Ministry of Health)*, 2010 BCCA 56, 2010 CarswellBC 275 (B.C. C.A.) at para. 27, leave to appeal refused 2010 CarswellBC 1851, 2010 CarswellBC 1852, [2010] 2 S.C.R. v (S.C.C.); *Gazankas v. Red Lake (Municipality)*, 2013 HRTO 198 (Ont. Human Rights Trib.) at para. 19; Ball, *supra* note 55, § 33:20.3; England, *Employment*

discriminatory *impact*, not discriminatory *intent*.[74] The Supreme Court has stated that the essence of discrimination in the employment context is "the arbitrariness of the barriers imposed, whether intentionally or unwittingly".[75] Accordingly, whether the workplace discrimination results from an employer's conscious decision to discriminate, a good faith decision that nonetheless has a discriminatory impact, a decision influenced by unconscious bias, or even a decision that arguably involves more automation than it does human judgment, it is all discrimination in the eyes of the law.

Just as there is no requirement to show intent to discriminate, there is no need to establish that a protected ground was the sole or even the main factor in the adverse impact; it need only be "a" factor.[76] Nor is there a requirement to show "directness".[77] In the hiring context, this proposition is expressly recognized in the *Canadian Human Rights Act*, which prohibits persons from "directly or indirectly" refusing to employ a person based on a prohibited ground,[78] and in the Ontario *Human Rights Code*, which contains a provision to the same effect.[79] In addition, the common law recognizes that prohibitions against indirect discrimination are implicit in human rights legislation, even where no explicit prohibition is present.[80] Indirect discrimination occurs "when discrimination is accomplished on prohibited grounds without explicit reference to the grounds".[81] The classic example is a facially neutral policy — e.g., "all

---

Law, *supra* note 46 at 214. As a corollary, a finding of discrimination does not necessarily entail a finding of moral blameworthiness. See *Action Travail des Femmes*, *supra* note 52 at 1134-35; Ball, *supra* note 55, § 33:20.3.

74 See Ball, *supra* note 55, § 33:10.4.

75 *MUHC*, *supra* note 53 at para 48.

76 See *Bombardier*, *supra* note 73 at para 52; *Moore v. British Columbia (Ministry of Education)*, 2012 SCC 61, 2012 CarswellBC 3446, 2012 CarswellBC 3447 (S.C.C.) at para. 33; *R. v. Bushnell Communications Ltd.*, 1974 CarswellOnt 470, 4 O.R. (2d) 288 (Ont. C.A.) at 290 [O.R.]; *Holden v. Canadian National Railway*, 1990 CarswellNat 847, 14 C.H.R.R. D/12 (Fed. C.A.) at D/15 [C.H.R.R.]; *Phipps v. Toronto Police Services Board*, 2010 ONSC 3884, 2010 CarswellOnt 7582 (Ont. Div. Ct.) at para. 76, affirmed 2012 ONCA 155, 2012 CarswellOnt 3992 (Ont. C.A.); *Rafiq v. Scotia Capital Inc.*, 2010 HRTO 697 (Ont. Human Rights Trib.) at para. 29.

77 See England, *Employment Law*, *supra* note 46 at 214. The distinction between direct and indirect discrimination, which used to be play a central role in determining the employer's duty of reasonable accommodation, was removed in *British Columbia (Public Service Employee Relations Commission) v. B.C.G.E.U.*, 1999 CarswellBC 1907, 1999 CarswellBC 1908, (*sub nom.* British Columbia (Public Service Employee Relations Commission) v. BCGSEU) [1999] 3 S.C.R. 3 (S.C.C.) [*Meiorin*]. See Stan Lanyon, "Conceptual Challenges in the Application of Discrimination Law in the Workplace" (2014) 3:1 Can. J. Hum. Rts. 75.

78 *Canadian Human Rights Act*, *supra* note 41, s. 7.

79 Ontario *Human Rights Code*, *supra* note 42, s. 9.

80 See *Simpsons-Sears*, *supra* note 60 at para 18; England, *Employment Law*, *supra* note 46 at 221-22.

employees are expected to work on Sundays" — that nonetheless has a discriminatory effect.[82]

With this legal framework in mind, the section below unpacks the ways in which hiring algorithms can mitigate certain hiring discrimination risks, as well as their limits in this respect and their potential capacity to introduce certain concerns over systemic discrimination.

III. HIRING ALGORITHMS AND WORKPLACE EQUALITY

(1) The Promise of Hiring Algorithms: Mitigating Certain Hiring Discrimination Risks

When used carefully, hiring algorithms may mitigate certain hiring discrimination risks. Support for this proposition begins with the recognition that human decisions, including those related to hiring, are influenced by the biases, stereotypes, and prejudices of their maker.[83] Of particular concern are unconscious biases, which operate outside conscious attentional focus. [84] The research on unconscious bias confirms that human decision makers "do not always have conscious, intentional control over the processes of social perception, impression formation, and judgment that motivate their actions".[85] Unconscious bias is of particular concern in the employment context because it is pervasive and plays a causal role in hiring discrimination.[86] As the Ontario Human Rights Tribunal stated in *Blakely v. Queen's University*,[87] "it is not uncommon that unstated and sometimes even unconscious biases may affect a hiring decision".[88]

Alongside unconscious bias is stereotyping.[89] According to social cognition theory, stereotyping is part of normal cognitive functioning.[90] This theory posits that our brains are wired to stereotype; it is something our brains do to "simplify the task of perceiving, processing, and retaining information about people in

---

81 Ball, *supra* note 55, § 33:20.2.

82 See *ibid*, § 33:20.6. See also *Québec (Commission des droits de la personne & des droits de la jeunesse) c. Gaz Métropolitain inc.*, 2011 QCCA 1201, 2011 CarswellQue 6683, 2011 CarswellQue 15856 (C.A. Que.) at para. 38.

83 See Kim, *supra* note 4 at 870.

84 See Anthony G Greenwald & Linda Hamilton Krieger, "Implicit Bias: Scientific Foundations" (2006) 94:4 Cal. L. Rev. 945 at 947.

85 See *ibid* at 946.

86 See *ibid* at 966.

87 *Blakely v. Queen's University*, 2012 HRTO 1177 (Ont. Human Rights Trib.).

88 *Ibid*. at para 40.

89 See Greenwald & Krieger, *supra* note 84 at 949 (defining a social "stereotype" as "a mental association between a social group or category and a trait").

90 See Linda Hamilton Krieger, "Content of Our Categories: A Cognitive Approach to Discrimination and Equal Employment Opportunity" (1995) 47:6 Stan. L. Rev. 1161 at 1188.

memory".[91] Once these stereotypes become ingrained, they influence intergroup judgment and decision making, "biasing in predictable ways the perception, interpretation, encoding, retention, and recall of information about other people", and they do so regardless of awareness or intention.[92] Though we may wish it were otherwise, we all engage in stereotyping and suffer from biases, whether conscious or unconscious.[93]

Turning to the hiring context, empirical studies performed in North America demonstrate that discrimination in hiring is not uncommon. In a study of data from a 2011 Canadian employment audit, researchers from Ryerson University and the University of Toronto found that for jobs requiring a university degree, Asian-named applicants had a 32.6 per cent lower interview selection rate than Anglo-named applicants, even when both groups had equivalent, all-Canadian qualifications.[94] In the United States, a well-known study published in 2000 found that when symphony orchestra auditions followed a blind process using a screen to conceal the candidate's identity from the jury, the chances of female musicians advancing beyond the preliminary round and ultimately being selected increased substantially.[95] In another study, researchers from MIT and the University of Chicago sent out 5,000 resumes in response to "help wanted" ads in Boston and Chicago.[96] They randomly assigned white-sounding names, such Emily or Greg, and African American-sounding names, such as Lakisha or Jamal, to the otherwise identical resumes. Their results, published in 2004, showed that applicants with white-sounding names received 50 per cent more call-backs than job applicants with African American-sounding names.

Research also demonstrates that we tend to hire those who are like "us" and to view "others" less favourably, with people tending to hire from their own social class, race, and gender.[97] Psychologically, this may be due to our need for

---

91 *Ibid.*

92 *Ibid.*

93 See Jerry Kang, "Trojan Horses of Race" (2005) 118:5 Harv. L. Rev. 1489 at 1496; Charles R. Lawrence II, "The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism" (1987) 39:2 Stan. L. Rev. 317 at 321. Both of the aforementioned articles are cited in Chander, *supra* note 4 at 1028, n 26.

94 Rupe Banerjee, Jeffrey G. Reitz & Phil Oreopoulos, "Do Large Employers Treat Racial Minorities More Fairly? An Analysis of Canadian Field Experiment Data" (2018) 44:1 Can. Pub. Pol'y 1.

95 Claudia Goldin & Cecilia Rouse, "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians" (2000) 90:4 Am. Econ. Rev. 715 at 716.

96 Marianne Bertrand & Sendhil Mullainathan, "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination" (2004) 94:4 Am. Econ. Rev. 991 at 997-99.

97 See Lauren A Rivera, "Hiring as Cultural Matching: The Case of Elite Professional Service Firms" (2012) 77:6 Am. Soc. Rev. 999; Joseph G. Altonji & Rebecca M. Blank, "Race and Gender in the Labor Market" in Orley Ashenfelter & David Card, eds., *Handbook of Labor Economics*, vol. 3C (Amsterdam: Elsevier, 1999) 3143, cited in CIHR, *Ethics of Algorithms*, *supra* note 11 at 5. "Ingroup favouritism" has been shown to

validation: by favouring those who we perceive to be like us, we validate our own self-worth.

Reliance on informal assessments by individual interviewers creates an increased risk of hiring discrimination.[98] As the Ontario Human Rights Commission describes, "conducting an interview by chatting with the applicant to see if he or she shares similar interests and will 'fit' into the organizational culture may present a barrier for persons who are or appear to be different than the dominant norm in the workplace. If this is used as a starting point for deciding whether candidates will be seen by senior decision-makers, this creates a major barrier to persons protected by the [*Human Rights Code*]".[99]

In sum, as the Centre for Internet and Human Rights observes, "human hiring systems are far from perfect".[100]

Enter hiring algorithms, which appear to offer an ostensibly objective and impartial means of selecting candidates that does not suffer from the cognitive imperfections and biases that we humans do. (More on that later.) In particular, they promise to remove conscious and unconscious biases from the hiring equation, resulting in a more objective and scientific process that better promotes workplace equality. In this way, they can be viewed as part of a broader movement in favour of evidence-based decision making.[101]

American law professor Anupam Chander argues that unconscious bias is less likely to manifest itself through the process of algorithm programming than through the process of human decision making because algorithm programming "requires a step-by-step writing process that depends on a conscious understanding of what is sought".[102] He adds that to the extent the discriminatory decisions humans make can be attributed to stereotypes formed through a process of statistical discrimination, algorithms acting on "richer information environments may not be subject to similar individually erroneous statistical discrimination".[103] In other words, because algorithms can absorb a more comprehensive dataset, they may be less susceptible to forming stereotypes, which in turn makes them better able to make non-discriminatory hiring decisions.

---

operate in the hiring context and result in discrimination. See Anthony G. Greenwald & Thomas F. Pettigrew, "With Malice Toward None and Charity for Some: Ingroup Favoritism Enables Discrimination" (2014) 69:7 American Psychologist 669 at 669; Dustin Volz, "Silicon Valley Thinks It Has the Answer to Its Diversity Problem", *The Atlantic* (26 September 2014), online: <www.theatlantic.com/politics/archive/2014/09/silicon-valley-thinks-it-has-the-answer-to-its-diversity-problem/431334>.

98 See Ontario Human Rights Commission, *supra* note 61.

99 *Ibid.*

100 CIHR, *Ethics of Algorithms*, *supra* note 11 at 5.

101 See World Wide Web Foundation, *supra* note 8 at 6.

102 Chander, *supra* note 4 at 1028-29.

103 *Ibid.* at 1030.

But hiring algorithms, like human decision makers, are far from perfect. Indeed, although they may address some of the flaws associated with human decision making, they may also introduce new, unanticipated issues.[104] Even leading AI recruitment startups recognize the risks: the founder and CEO of one such startup acknowledged in an interview that there is a "huge risk that using AI in the recruiting process is going to increase bias and not reduce it".[105] It is therefore essential to explore the limitations on hiring algorithms and to examine their potential capacity to introduce concerns over discrimination, and in particular systemic discrimination.

(2) The Challenges: Limitations and Concerns over Systemic Discrimination

(a) Bad Data, Bad Rules, or Both

Hiring algorithms may fail to deliver on their promise to the extent that they act on bad data, apply bad rules, or do both.

i. Bad Data

The data fed into hiring algorithms will inevitably be tainted by real-world biases, stereotypes, and injustices, and this may lead to a discriminatory outcome that reinforces and reproduces marginalization.[106] This is particularly so for machine learning algorithms. As Chander writes, "[e]ven facially neutral algorithms will produce discriminatory results because they train and operate on the real world of pervasive discrimination".[107] English researchers Bryce Goodman and Seth Flaxman make a similar observation: "machine learning can reify existing patterns of discrimination — if they are found in the training dataset, then by design an accurate classifier will reproduce them. In this way, biased decisions are presented as the outcome of an 'objective' algorithm."[108] Scholars Solon Barocas and Andrew Selbst put it this way: "an algorithm is only as good as the data it works with".[109] As the saying goes, "garbage in, garbage out".[110]

---

104 See CIHR, *Ethics of Algorithms*, *supra* note 11 at 6.

105 Chandler, *supra* note 31.

106 See World Wide Web Foundation, *supra* note 8 at 9; Faisal Kamiran & Toon Calders, "Classifying Without Discriminating" (Paper delivered at the Second International Conference on Computer, Control and Communication, Karachi, Pakistan, 17-18 February 2009) at 1, online: <ieeexplore.ieee.org/document/4909197/>, cited in Chander, *supra* note 4 at 1036.

107 Chander, *supra* note 4 at 1036. See also *ibid* (claiming that "ostensibly neutral algorithms can produce results that reflect the prejudices of society"); Moritz Hardt, "How Big Data Is Unfair", *Medium* (26 September 2014), online: <medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de>; Joh, *supra* note 27 at 300.

108 Bryce Goodman & Seth Flaxman, "European Union Regulations on Algorithmic Decision-Making and a 'Right to Explanation'" (2017) 38:3 A.I. Magazine 50 at 3, online: Cornell University Library <arxiv.org/abs/1606.08813>. See also Kim, *supra* note 4 at 861.

It is not hard to imagine how this issue can arise in the hiring algorithm context. Consider, for example, an industry that has come to be male-dominated as a result of years of systemic discrimination against women. If a hiring algorithm is designed (or learns) to assign weight to years of industry experience, then women will tend to score less favourably than men simply because they have historically been shut out of the industry. Thus, by acting on data tainted by discrimination, the hiring algorithm reproduces historical injustices.

Barocas and Selbst highlight a further data-related issue: disadvantaged groups may be misrepresented in datasets upon which algorithms rely, producing discriminatory results.[111] The researchers cite concerns over "the nonrandom, systemic omission of people who live on big data's margins, whether due to poverty, geography, or lifestyle, and whose lives are less 'datafied' than the general population's",[112] as well as the concern that "[b]ecause not all data is created or even collected equally, there are 'signal problems' in big-data sets — dark zones or shadows where some citizens and communities are overlooked or underrepresented".[113] They suggest that as a consequence of the higher incidence of data misrepresentation in the case of members of disadvantaged groups, the conclusions drawn from the data will be skewed, which may have a disproportionate and discriminatory effect on protected groups to the extent that the group's disadvantage is correlated with protected characteristics.[114]

We can readily envision how this type of data misrepresentation and resulting discrimination might occur in the hiring algorithm context. For example, members of marginalized communities with limited access to the Internet have an impaired ability to build up their professional presence on the Internet (e.g., a LinkedIn profile, a blog, a personal webpage displaying their personal portfolio, etc.). If a hiring algorithm is designed (or learns) to assign weight to the quality of a candidate's Internet presence, members of marginalized communities with limited access to the Internet would be at a distinct disadvantage.

One group of researchers suggests a potential response to the issue of bad data: partially repairing the data fed into the algorithm so as to make the data

---

109 Barocas & Selbst, *supra* note 4 at 671. See also Chandler, *supra* note 31; Jackie Snow, "New Research Aims to Solve the Problem of AI Bias in 'Black Box' Algorithms", *MIT Technology Review* (7 November 2017), online: <www.technologyreview.com/s/609338/new-research-aims-to-solve-the-problem-of-ai-bias-in-black-box-algorithms>; Chander, *supra* note 4 at 1036.

110 Barocas & Selbst, *supra* note 4 at 683.

111 *Ibid.* at 684-85.

112 *Ibid.* at 684, citing Jonas Lerman, "Big Data and Its Exclusions" (2013) 66 Stan. L. Rev. Online 55 at 57, online: <www.stanfordlawreview.org/online/privacy-and-big-data-big-data-and-its-exclusions>.

113 Kate Crawford, "Think Again: Big Data", *Foreign Policy* (May 10, 2013), online: <www.foreignpolicy.com/articies/2013/05/09/thinkagain-bigdata>.

114 Barocas & Selbst, *supra* note 4 at 684.

unbiased.[115] While this process may mitigate the issue to some extent, it is an incomplete solution. One can repair biased data only so much, and identifying when and in what ways data is biased may be challenging if not impossible in some cases.

ii. Bad Rules

Beyond the issue of bad data, a hiring algorithm based on poorly designed rules is liable to produce discriminatory results. Obviously, an algorithm that is deliberately designed to disfavour certain candidates based on protected characteristics is discriminatory, but in practice bad rules ares more likely to arise inadvertently. For example, consider a hiring algorithm that is designed (or learns) to favour candidates who have steady, uninterrupted work histories — this may be seen as a proxy for dependability. Such a rule would disadvantage women since, in Canada, women are statistically more likely than men are to take parental leave or to temporarily withdraw from the workforce in order to raise children.[116] In legal terms, the outcome may be systemic discrimination: "policies or procedures [that] impact disproportionately on certain statutorily protected groups".[117] Importantly, based on the legal principles identified in Part II of this article, this conclusion holds whether or not the rule was intended to produce discriminatory results. What matters is discriminatory *impact*, not discriminatory *intent*.

The issue of bad rules becomes more complex where machine learning algorithms are involved. As computer scientist Ben Shneiderman observes, machine learning algorithms are "more concerning because if you're automating a process, then you're reducing the opportunities for a human being to check the bias".[118] The machine learning algorithm may teach itself to apply discriminatory rules, and the opportunity for human intervention is reduced. For example, it may determine that there is a correlation between race and promotion rates and develop a discriminatory hiring rule based on this correlation, without any direction from its human creator.

As this example demonstrates, machine learning algorithms may discover and act on correlations that have little or no connection to the candidate's actual qualifications. The data mining techniques employed by machine learning algorithms generally seek out "any statistical relationship between variables present in the data, regardless of whether the reasons for the relationship are understood".[119] This can lead to problems. For example, based on past data, a

---

115 Feldman et al, "Certifying and Removing Disparate Impact" (Extended version of paper accepted at 2015 ACM SIGKDD Conference on Knowledge Discovery and Data Mining), online: Cornell University Library <arxiv.org/abs/1412.3756>.

116 See Ontario Human Rights Commission, *supra* note 61.

117 *Crockford*, *supra* note 68 at para 49.

118 Christina Couch, "Ghosts in the Machine", *PBS* (25 October 2017), online: <www.pbs.org/wgbh/nova/next/tech/ai-bias>.

119 Kim, *supra* note 4 at 865.

machine learning algorithm may discover that the name "Rick" is associated with higher work performance. The discriminatory impact that a decision-making rule based on this correlation would have been obvious. Or a machine learning algorithm may learn that people who live closer to work are statistically more likely to stay with the company over the long term, and it may in turn develop a hiring rule based on this correlation. Yet this is another "bad rule". One's area code has no bearing on one's qualifications. This is problematic because, as the Ontario Human Rights Commission explains, "Employers must make sure that only information about qualifications and job requirements is considered when making hiring decisions".[120] Consequently, prospective employers who rely on hiring algorithms risk denying candidates positions based on unexplained correlations that have little to no connection to the candidate's actual qualifications.[121] Furthermore, the rule developed by the algorithm in the example above may have a disparate impact on protected groups to the extent that race, nationality, and other protected characteristics may be correlated with residency. Similar discriminatory effects may arise through other correlations discovered and acted upon.

However, the risk of bad rules can be mitigated to some degree through careful study, design, and monitoring. Machine learning algorithms should not simply be left to their own devices. Shneiderman suggests that algorithm logs that are capable of comprehension should be produced so that the basic rules underlying an algorithm can be kept in check.[122] That said, while the risk of algorithmic discrimination can be mitigated to some degree, it cannot be excluded entirely.[123]

Finally, while it may seem like an attractive solution to simply remove the consideration of protected characteristics from the hiring algorithm's calculus, this will not solve the problem, and in fact several scholars have argued persuasively that omitting protected characteristics would be counterproductive and would only increase the risk of discrimination.[124] The simple removal of a variable from a dataset does not mean it will not appear elsewhere in the data, as other variables that have a close correlation with the excluded variable may serve as proxies.[125] As U.S. law professor Pauline Kim writes, "[b]ecause other information contained in large datasets can serve as a proxy for race, disability, or other protected statuses, simply eliminating data on those characteristics cannot prevent models that are biased along these dimensions".[126]

---

120 Ontario Human Rights Commission, *supra* note 61.

121 See Kim, *supra* note 4 at 865; King & Mrkonich, *supra* note 4 at 572-73.

122 See Couch, *supra* note 118; Ben Shneiderman, "The Dangers of Faulty, Biased, or Malicious Algorithms Requires Independent Oversight" (2016) 113: 48 P.N.A.S. 13538, online: <www.pnas.org/content/pnas/113/48/13538.full.pdf>.

123 See Volz, *supra* note 97.

124 See e.g. Joshua A. Kroll et al., "Accountable Algorithms" (2017) 165:3 U. Penn. L. Rev. 633 at 685; Kim, *supra* note 4 at 904, 918; Chander, *supra* note 4.

125 See Kim, *supra* note 4 at 880, citing Barocas & Selbst, *supra* note 4.

iii. Both

The issues of bad data and bad rules may overlap — indeed, they may be inseparable. As explained above, algorithms must be viewed in their broader context, as they form part of an interconnected system.[127] Biases, stereotypes, and injustices may be traced to the designer of the algorithm, the data to which it is applied, society more broadly, or some combination of the foregoing.[128] The interaction between bad rules and bad data may create a feedback loop that perpetuates and exacerbates existing biases, stereotypes, and injustices.

Finally, whether the problem is bad data, bad rules, or both, the concerns expressed above cannot be dismissed on the basis that the blame lies not with humans, but with hiring algorithms. From a legal perspective, no existing doctrine permits an employer to avoid liability for hiring discrimination on the basis that the discriminatory effect stemmed from its decision to use a particular technology, though in theory it might be able to seek an indemnity from an algorithm supplier if permitted by the parties' contract. On a more philosophical level, all hiring algorithms, even those that use machine learning technology, are designed and implemented by humans. As such, we must take responsibility for the results produced by hiring algorithms. This sentiment is captured by the Fairness, Accountability, and Transparency in Machine Learning group, which observes that "[a]lgorithms and the data that drive them are designed and created by people — [t]here is always a human ultimately responsible for decisions made or informed by an algorithm . . . '[t]he algorithm did it' is not an acceptable excuse".[129]

There is a further risk that the problems identified above will be overlooked based on a false belief that hiring algorithms are inherently fair and objective. That risk is discussed below.

(b) Ostensible Fairness and Objectivity

Hiring algorithms appear to be inherently fair and objective. As Chander writes, "[a]lgorithms can make decisionmaking seem fair precisely because computers are logical entities which should not be infected by all-too-human bias".[130] The use of algorithms may thereby give the decision-making process "'a patina of inevitability' . . . and indeed a patina of fairness".[131] Similarly, as U.S.

---

126 Kim, *supra* note 4 at 898.

127 See World Wide Web Foundation, *supra* note 8 at 9.

128 See *ibid*, citing Kate Crawford & Jason Schultz, "Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms" (2014) 55:1 B.C. L. Rev. 93; Barocas & Selbst, *supra* note 4.

129 Fairness, Accountability and Transparency in Machine Learning, "Principles for Accountable Algorithms", online: <www.fatml.org/resources/principles-for-accountable-algorithms>.

130 Chander, *supra* note 4 at 1034. See also Hardt, *supra* note 107.

131 Chander, *supra* note 4 at 1034, citing Frank Pasquale, *The Black Box Society: The Secret*

law professor Elizabeth Joh writes, "the very idea of algorithmic decisionmaking may subtly appeal to us as objective because of its mathematical basis".[132] Data scientist Fred Benenson cautions against giving in to "mathwashing", the assumption that algorithmic models are free from subjectivity because they involve math.[133] There is a sense that because the algorithm is acting on *data*, not merely on personal impressions, its output must be fair. There is also a sense of inevitability that attaches to the outcome of an algorithm — a sense that the result could not have been any other way. This lends legitimacy to hiring algorithms.

But despite the apparent fairness and objectivity of hiring algorithms, the results they produce may nonetheless be discriminatory. Just as an objective, facially neutral policy subjecting different groups to the same standards can result in *prima facie* discrimination,[134] so too can the application of an objective, facially neutral rule applied by a hiring algorithm.[135] The Canadian jurisprudence makes clear that the mere fact that an employment policy is applied uniformly to all applicants, and is used for sound business reasons in good faith, does not mean its application cannot result in unlawful discrimination.[136] This principle can be applied in the context of hiring algorithms: a hiring algorithm may be designed with the best of intentions and may be equally applicable to all job applicants, but if it produces a discriminatory result, these features will provide no defence.

Moreover, scholars have challenged the notion that algorithms are objective.[137] Researcher Tarleton Gillespie, describes the "promise of algorithmic objectivity" as "the way the technical character of the algorithm is positioned as an assurance of impartiality".[138] He suggests that the performance of "algorithmic objectivity" is fundamental to the legitimacy of algorithms.[139] He writes: "More than mere tools, algorithms are also stabilizers of trust, practical and symbolic assurances that their evaluations are fair and accurate, and free from subjectivity, error, or attempted influence. But, though algorithms may

---

Algorithms That Control Money and Information (Cambridge, Mass.: Harvard University Press, 2015) at 15.

132 Joh, *supra* note 27 at 292.

133 Tyler Woods, "'Mathwashing,' Facebook and the Zeitgeist of Data Worship", *Technical.ly*, online: <technical.ly/brooklyn/2016/06/08/fred-benenson-mathwashing-facebook-data-worship>, cited in Joh, *supra* note 27 at 292.

134 See *Meiorin*, *supra* note 77.

135 See Chander, *supra* note 4 at 1024.

136 See *Simpsons-Sears*, *supra* note 60 at para 18.

137 See, e.g., Gideon Mann & Cathy O'Neil, "Hiring Algorithms Are Not Neutral", *Harvard Business Review* (9 December 2016), online: <hbr.org/2016/12/hiring-algorithms-are-not-neutral>.

138 Gillespie, *supra* note 6 at 168.

139 *Ibid.* at 180.

appear to be automatic and untarnished by the interventions of their providers, this is a carefully crafted fiction."[140]

Scholars have also observed that although algorithms may be assumed to be objective and free from bias, an algorithm "will reflect the perspective and biases of its creators".[141] Gideon Mann and Cathy O'Neil argue that "[a]lgorithms are, in part, our opinions embedded in code. They reflect human biases and prejudices that lead to machine learning mistakes and misinterpretations".[142]

In sum, despite their appearance, hiring algorithms are not inherently fair and objective. But even if users recognize the need to critically assess hiring algorithms despite their apparent fairness and objectivity, they may be limited in their ability to do so as a result of the proprietary nature and opacity of many algorithms.

(c) Proprietary Nature and Opacity

In a June 2017 white paper entitled *Algorithmic Accountability*,[143] the World Wide Web Foundation acknowledged both the growing need to scrutinize algorithms and the barriers to performing that scrutiny, emphasizing that "the outcomes of algorithmic processes are often not designed to be accessible, verified or evaluated by humans, limiting our ability to identify if, when, where, and why the algorithm produced harm — and worse still — redress this harm".[144] Consistent with observation, algorithms have frequently been described in the literature as "black boxes"[145] and have been criticized as being both "proprietary and opaque".[146] Joh suggests that the term "black box" can be applied to algorithms in two senses. First, "the calculations used [by the algorithm] to make a decision may be inscrutable to the person affected by that

---

[140] *Ibid.* at 179.

[141] Joh, *supra* note 27 at 292.

[142] Mann & O'Neil, *supra* note 137.

[143] World Wide Web Foundation, *supra* note 8.

[144] World Wide Web Foundation, *supra* note 8 at 4.

[145] See e.g. Pasquale, *supra* note 131 at 3 (Pasquale describes a "black box" as "a system whose workings are mysterious; we can observe its inputs and outputs, but we cannot tell how one becomes the other"); Nicholas Diakopoulos, *Algorithmic Accountability Reporting: On the Investigation of Black Boxes* (Tow Center for Digital Journalism, Columbia Journalism School, December 2013), online: <towcenter.org/wp-content/uploads/2014/02/78524_Tow-Center-Report-WEB-1.pdf>; World Wide Web Foundation, *supra* note 8 at 12; Claire Cain Miller, "When Algorithms Discriminate", *The New York Times* (9 July 2015), online: <www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html>. But see Gillespie, *supra* note 6 at 178 (suggesting that the "black box" metaphor is inapt to the extent that algorithms may be adjusted or "tweaked" by its creator).

[146] Sarah Tan et al., "Detecting Bias in Black-Box Models Using Transparent Model Distillation" (Paper delivered at the Neural Information Processing Systems 2017 Symposium on Interpretable Machine Learning, Montreal, Canada, updated 18 November 2017) at 1, online: <arxiv.org/abs/1710.06169>.

decision", and "as machine learning algorithms become more complex, they may be inscrutable to the programmers themselves".[147] Second, "the companies that create [algorithms] often refuse to divulge information about them".[148]

Hiring algorithms may be inaccessible by design. Suppliers may offer clients the right to use their algorithms for a fee, without revealing the underlying code, which is protected as proprietary. It has been suggested that most employers who use hiring algorithms rely on "out-of-the-box" algorithms without fully understanding how those algorithms actually work.[149] In addition, some algorithms are subject to non-disclosure agreements.[150] Furthermore, it has been suggested that while hiring algorithms could be made open source, it would be unrealistic to insist on complete transparency, as doing so would result in public disclosure of commercial proprietary information and trade secrets.[151]

The proprietary nature and opacity of hiring algorithms limit the ability to critically assess the algorithm and identify whether (and if so, why) it may create or perpetuate injustices, still less resolve the issue.[152] In the words of the Centre for Internet and Human Rights, "[i]f an algorithm is opaque, it becomes impossible for outsiders to understand the rationale behind any particular outcome, or when algorithms are misused".[153] Put differently, it is impossible to take a look under the hood.

Machine learning algorithms present a higher level of opacity than "plain vanilla" algorithms. Machine learning algorithms do not operate on the basis of factors and weights identified in advance by its human programmers; rather, they construct a model based on the data correlations it discovers in the data to which it is exposed.[154] As Kim writes, "[w]hen [a machine learning algorithm] is relied on to screen or rank applicants, it obscures the basis on which employers are making ultimate employment decisions. This lack of transparency makes it difficult to know if any observed bias is simply a byproduct of justifiable business considerations or the result of flaws in the model's construction".[155] Moreover, a machine learning algorithm may reshape itself over time to the point where we can no longer understand it.[156]

One consequence of the opaque, proprietary nature of hiring algorithms is that it becomes very difficult for anyone, including potential claimants, to detect

---

147 Joh, *supra* note 27 at 292-93.

148 *Ibid.* at 293.

149 See World Wide Web Foundation, *supra* note 8 at 12.

150 See Christina Counch, "Ghost in the Machine", *PBS* (25 October 2017), online: <www.pbs.org/wgbh/nova/next/tech/ai-bias/>.

151 See Kroll et al, *supra* note 124 at 639, 658, cited in Chander, *supra* note 4 at 1040.

152 World Wide Web Foundation, *supra* note 8 at 4.

153 CIHR, *Ethics of Algorithms*, *supra* note 11 at 11.

154 Kim, *supra* note 4 at 881.

155 *Ibid.* at 881.

156 See Chander, *supra* note 4 at 1040; CIHR, *Ethics of Algorithms*, *supra* note 11 at 3.

whether unlawful discrimination may have occurred, and even more difficult to find proof. The Canadian case law recognizes that in the absence of a "smoking gun", it is exceedingly difficult to establish hiring discrimination.[157] That difficulty is particularly acute in the context of hiring algorithms. A hiring algorithm that takes in massive amounts of data and applies a complex or even unknown formula to make a hiring decision may make it practically impossible to determine whether discrimination has occurred.

The opacity of hiring algorithms can present legal issues for employers as well. For example, if a *prima facie* case of discrimination is made out, the opacity of hiring algorithms may present an impediment to the employer's ability to discharge its onus of offering a reasonable, non-discriminatory explanation for why the claimant was denied employment. Accordingly, both prospective employees and the prospective employers have a shared interest in ensuring hiring algorithms are transparent and understandable. However, this shared interest sits at tension with the interest of the supplier of the hiring algorithm, which is to ensure its source code remains a proprietary trade secret.

One potential response to the opacity of hiring algorithms is to require transparency through legislation. In response to concerns over the black box nature of algorithms, some have advocated for a "right to explanation" (i.e., an individual's right to an explanation of how algorithmic decisions affecting them are made). Indeed, since it came into force in May 2018, the European Union's General Data Protection Regulation[158] (GDPR) has, at least in effect, recognized a right to explanation (in certain circumstances) across all EU member states.[159] Articles 13(2)(f), 14(2)(g), and 15(1)(h) of the GDPR require "data controllers" to provide "data subjects", neither of which is defined in the GDPR, with information about "the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject". Article 22(1) provides that data subjects "have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her". Article 22(2) identifies circumstances in which decisions based solely on automated processing are permitted, such as where the decision "(a) is necessary for entering into, or

[157] See *Basi*, *supra* note 65 at D/5038.

[158] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC [2016] OJ L 119/1.

[159] See Andrew D. Selbst & Julia Powles, "Meaningful Information and the Right to Explanation" (2017) 7:4 Int'L Data Privacy Law 233; Goodman & Flaxman, *supra* note 108. But see Sandra Wachter, Brent Mittelstadt & Luciano Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation" (2017) 7:2 Int'l Data Privacy Law 76 (arguing that the GDPR does not create a "right to explanation").

performance of, a contract between the data subject and a data controller" or "(c) is based on the data subject's explicit consent". Article 22(3) provides that where either of these two conditions apply, "the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision." Article 22(4) stipulates that decisions made under Article 22(2) shall not be based on special categories of personal data[160] unless certain conditions are met. These provisions could readily be applied in the context of hiring algorithms, and they offer one model that could be adopted through domestic legislation in Canada.

Another response would be to establish one or more public bodies to regulate and oversee algorithms. The idea has already been floated in the United States. For example, Andrew Tutt, a lawyer based in Washington, D.C., has suggested that a specialist regulatory agency be formed to oversee the use and sale algorithms.[161] He envisions an agency similar to the U.S. Food and Drug Administration. Essentially, under Tutt's model, hiring algorithms could be treated as a potentially hazardous product, subject to regulatory oversight and testing. Similarly, American computer scientist and professor Ben Shneiderman has proposed an "Algorithm Safety Board" that would oversee high-stakes algorithms and investigate problems.[162]

But even if algorithms were transparent in their design, several researchers have argued that this would still not be enough.[163] For example, Mike Ananny and Kate Crawford argue that being able to see a system is insufficient to understand how that system actually works and to govern it.[164] Similarly, Cynthia Dwork and Professor Deirdre Mulligan maintain that "it is unreasonable to expect transparency alone to root out bias".[165] Chander argues that rather than transparency in the *design* of the algorithm, what is needed is transparency in its *inputs and outputs*.[166] He proposes that the focus be

---

160 Those special categories are listed in Article 9(1): "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation".

161 Andrew Tutt, "An FDA for Algorithms" (2016) 69 Admin. L. Rev. 83.

162 Ben Shneiderman, "The Dangers of Faulty, Biased, or Malicious Algorithms Requires Independent Oversight" (2016) 113: 48 P.N.A.S. 13538, online: <www.pnas.org/content/pnas/113/48/13538.full.pdf>.

163 See generally Kroll et al, *supra* note 124.

164 Mike Ananny & Kate Crawford, "Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability" (2016) New Media & Society 1 at 1, cited in World Wide Web Foundation, *supra* note 8 at 10.

165 Cynthia Dwork & Deirdre K. Mulligan, "It's Not Privacy, and It's Not Fair" (2013) 66 Stan. L. Rev. Online 35 at 37, online: <review.law.stanford.edu/wp-content/uploads/sites/3/2016/08/DworkMullliganSLR.pdf>, cited in Chander, *supra* note 4 at 1040.

166 Chander, *supra* note 4 at 1039.

placed not on merely seeking transparency of algorithmic design, but rather on the following question: Are protected groups receiving statistically worse results, given their relevant characteristics?[167] Chander thus places the focus on outcomes, rather than on the mechanics of the algorithm.

Finally, although hiring algorithms may be criticized as "black boxes", the human mind is no more transparent.[168] As Chander argues, "[t]he ultimate black box is the human mind" and "[p]rejudices acted upon in this black box never have to be written down".[169] For this reason, it cannot be said that relying solely on human decision making would remove all concerns over opacity in the hiring process. There will always be some level of opacity in hiring decisions. But that does not mean we should take steps to make hiring algorithms, in both their design and their inputs and outputs, more transparent and intelligible.

IV. CONCLUSION

A number of commentators have expressed optimism about the potential social and legal impacts of algorithms. [170] In the employment context, there is good reason to think that hiring algorithms can, if used carefully, mitigate certain discrimination risks. The overarching message of this article is not that algorithms have no place in the hiring process or that they cannot assist in achieving greater workplace equality. Rather, it is that they are no panacea against hiring discrimination. As long as biases, prejudices, and injustices exist, hiring algorithms will run the risk of reproducing and reinforcing them. Hiring algorithms cannot eliminate discrimination, and in fact they may introduce certain concerns over systemic discrimination. Accordingly, we should view hiring algorithms for what they are: a set of rules designed to solve a particular problem. These rules are not perfect, nor are the data they act upon. If employers

---

167 *Ibid.*

168 See Kroll et al, *supra* note 124 at 634 (observing that "[t]he implicit (or explicit) biases of human decisionmakers can be difficult to find and root out, but we can peer into the 'brain' of an algorithm").

169 *Ibid.* at 1030.

170 See e.g. Goodman & Flaxman, *supra* note 108 at 7 (arguing that "properly applied, algorithms can not only make more accurate predictions, but offer increased transparency and fairness over their human counterparts"); Rónán Kennedy, "Algorithms and the Rule of Law" (2017) 17 Legal Info. Mgmt. 170 at 172 (stating that "[d]igital technology . . . offers opportunities for transparency and empowerment, and properly designed systems may . . . help to overcome bias and prejudice."); Kim, *supra* note 4 at 873 (maintaining that "[d]ata analytics . . . hold the potential to reduce biases and increase opportunities in the workplace for traditionally disadvantaged groups. But much depends on how data are used"); Florentine, *supra* note 29 (citing the CEO of an industry advocacy group who claims that "[w]hether by eliminating unconscious bias in general or attacking specific manifestations of bias in recruiting, screening and hiring talent, AI and machine learning has potential to level the playing field for women and other underrepresented minorities and provide a competitive advantage for companies").

decide to use hiring algorithms, they should do so in a careful, conscientious manner, fully aware of the benefits they offer and the risks they present.