# Busting the Black Box: Big Data, Employment and Privacy

By: **Rebecca J. Wilson, Kiley M. Belliveau and Leigh Ellen Gray**

IADC member, Rebecca J. Wilson, is a partner in the Boston office of Peabody & Arnold where she concentrates her practice in the representation of employers in employment litigation and employment practices counseling. Ms. Wilson has been selected for inclusion in Best Lawyers in America for her expertise in Employment Law-Management and Litigation-Labor and Employment. Ms. Wilson is a past member of the IADC Board of Directors and Past Chair of its Employment Law Committee.

Kiley M. Belliveau is a partner in the Boston office of Peabody & Arnold LLP and a 2007 graduate of the IADC Trial Academy. Ms. Belliveau defends employers and managers in all phases of civil litigation. In addition to her employment litigation practice, she enjoys working with clients to address the myriad compliance, training, and risk management issues that can arise out of the employment relationship.
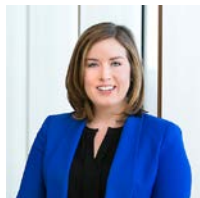
Leigh Ellen Gray is an associate in the Boston office of Peabody & Arnold and a 2014 graduate of the Charleston School of Law, where she served as Editor-in-Chief of the Charleston Law Review. Before relocating to Boston, Ms. Gray clerked for the Honorable P. Michael Duffy, Senior United States District Court Judge for the District of South Carolina and the Honorable Carmen T. Mullen, South Carolina Circuit Court Judge. Ms. Gray practices in the areas of employment litigation and counseling, medical malpractice, premises liability and tort defense.

WE live in an era of big data. Our increasing reliance on digital communication coupled with the technological ability to capture, collect, and analyze ever-growing volumes of data has led to the application of predictive analytics techniques to many of the most important facets of our lives, including healthcare, education, and employment.[1] The ubiquitous nature of big data raises questions about "the relationship between individuals and those who collect and use data about them."[2] In the seminal Harvard Law Review article "The Right to Privacy," Samuel D. Warren and Louis D. Brandeis wrote of the need for the law to adapt to address new intrusions on the right to privacy occasioned by social and technological change. The article opens with the following observation:

> That the individual shall have full protection in person and in property is a

principle as old as the common law; but it has been found necessary from time to time to define anew the exact nature and extent of such protection. Political, social, and economic changes entail the recognition of new rights, and the common law, in its eternal youth, grows to meet the demands of society.[3]

These prescient words apply as forcefully today in the age of big data as they did in 1890 when Warren and Brandeis first discussed "the right to be let alone."[4] Inherent in the traditional view of the right to privacy is the generally accepted principle that each individual has "the right of determining, ordinarily, to what extent his thoughts, sentiments, and emotions shall be communicated to others."[5] Yet in our wired world, individuals passively communicate infor-

---

[1] *See, e.g.*, EXEC. OFFICE OF THE PRESIDENT, BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES (2014), [hereinafter EOP MAY 2014 REPORT], available at http://www.white house.gov/sites/default/files/docs/big_dat a_privacy_report_5.1.14_final_print.pdf, *archived at* http://perma.cc/9PNW-JBEK.
[2] *Id.* at 3.
[3] Samuel D. Warren and Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 193 (1890).

[4] *Id.* at 193-194.
[5] *Id.* at 198, citing Millar v. Taylor, 4 Burr. 2303, 2379 (1769) ("It is certain every man has a right to keep his own sentiments, if he pleases. He has certainly a right to judge whether he will make them public, or commit them only to the sight of his friends.").

mation about themselves each day with little knowledge about or control over how the information is transmitted and the purposes for which it is used. Big data raises concerns about not only the individual right to privacy, but also whether it creates "such an opaque decision-making environment that individual autonomy is lost in an impenetrable set of algorithms."[6] Existing legal frameworks may prove insufficient to address novel privacy concerns raised by big data, and the time might yet again be upon us to consider the scope of the right to privacy and the legal mechanisms required to protect it. This article focuses on the use of big data in the employment context. Big data can be used by employers in many positive ways, including eliminating irrational or even discriminatory biases in the hiring process; identifying unique and unexpected sources of talent; promoting employee wellness; reducing healthcare costs; and increasing worker efficiency. Critics of predictive analytics in the workplace decry the fact that data from digital activities can be used by employers to make assumptions about individuals' behavior that impact their livelihood without their even knowing it.[7] Employers

and their advisors who seek to realize the potential of big data must navigate largely uncharted territory because big data does not fit neatly within existing legal frameworks that govern the employment relationship. Until employment laws are updated to more directly address big data, counsel advising employers on the use of big data in the workplace must consider how existing legal protections may apply. Many compliance issues can arise and will continue to arise as the technology evolves and new applications emerge. This article seeks to provide employers and their counsel with just a few examples of the impact that big data can have in the workplace and the related compliance concerns.

I. Defining Big Data

A. Characteristics of Big Data

There are many different definitions of big data. In the privacy context, big data has been defined as "data about one or a group of individuals, or that might be analyzed to make inferences about individuals."[8] Perhaps the

---

[6] EOP MAY 2014 REPORT, *supra* note 1, at 10.
[7] Ian Kerr and Jessica Earle, *Prediction, Preemption, Presumption: How Big Data Threatens Big Picture Privacy*, 66 STAN. L. REV. ONLINE 65, 71 (2013) ("big data can be

used to make important decisions that implicate us without our even knowing it").
[8] EXEC. OFFICE OF THE PRESIDENT, BIG DATA AND PRIVACY: A TECHNOLOGICAL PERSPECTIVE 2 (2014), [hereinafter EOP BIG DATA AND

most commonly-referenced characteristics that make data "big" are the so-called "three V's:" datasets of enormous *volume*, in an ever-increasing *variety* of formats, continuously collected at a rapid *velocity*. [9] This framework recognizes that, first, routine data collection is now deeply embedded in many aspects of our daily lives, and, second, these datasets are ripe for computer-assisted or automated analysis. [10] As scholar and big data ethicist Dr. Solon Barocas puts it, "the distinguishing feature of big data [is] the ability to detect useful patterns in datasets that can inform or automate future decision making…data is big when it can function as the grist for the analytics mill." [11]

1. The "Three Vs"

A. Volume

It can be difficult to comprehend the volume of data created and shared in today's hyper-connected world. For example, it is estimated that in 2016

the amount of data transferred, for the first time, crossed the one zettabyte threshold. [12] If you consider that a byte of information translates to one character of text, Tolstoy's *War and Peace*, which clocks in at 1,250 pages, would fit into a zettabyte 323 trillion times. [13] Another comparator: consider that if every person in the United States took a digital photo every second of every day for over a month, all of those photos put together would equal roughly one zettabyte. [14] But for as much data as people create—for example, an average of 500 million photos per day and over 200 hours of video per minute shared in 2014—that volume is nothing compared with the amount of digital information created *about* them each day. [15]

B. Variety

Big data is varied. It is generated in many different forms, and it is captured and transmitted via an array of applications. Big data sources can be divided into two basic categories: data that is "born

PRIVACY], available at https://obama whitehouse.archives.gov/sites/default/file s/microsites/ostp/PCAST/pcast_big_data_ and_privacy_-_may_2014.pdf.

[9] *See, e.g.*, EOP MAY 2014 REPORT, *supra* note 1, at 4.

[10] *EEOC at 50: Progress and Continuing Challenges in Eradicating Employment Discrimination*, Meeting of the Equal Employment Opportunity Commission (Jul. 1, 2015) [hereinafter *EEOC at 50 July Meeting*] (statement of Dr. Solon Barocas,

Postdoctoral Research Associate Center for Information Technology Policy, Princeton University), available at https://www.eeoc. gov/eeoc/meetings/7-1-15/barocas.cfm.

[11] *Id.*

[12] Tim Willingham, *2016: The Year of the Zettabyte*, DAILY INFOGRAPHIC (Mar. 23, 2013), http://dailyinfographic.com/2016-the-year-of-the-zettabyte-infographic.

[13] *Id.*

[14] EOP MAY 2014 REPORT, *supra* note 1, at 2.

[15] *Id.*

digital," meaning that it is specifically created for use by a computer or data processing system, and data that is "born analog," meaning that it originates in the tangible, physical world but can be converted into digital data.[16] Examples of data that is "born digital" include data: contained in emails (including content, frequency, recipients, and read receipts); generated from web browsing; captured by items that make up the Internet of things ("smart" devices such as digital assistants like the Amazon Echo, wearable fitness monitors, or Internet-connected cars); collected through store loyalty programs which track your purchases online and in stores; about your location, gathered from GPS, cell tower triangulation, wireless network utilization, and card swipe security systems; generated and shared on social media; collected from mobile applications; and, in the context of employment, generated by performance on psychometric

tests.[17] Some examples of data that is "born analog" but can then be digitized include sound waves in phone calls, content from video footage, and documents that are scanned and run through optical character recognition (OCR) software.[18]

C. Velocity

The "velocity" of big data refers to both the swift pace of data collection [19] as well as the continuity of the data stream.[20] For example, mobile mapping applications are useless unless they are constantly harvesting the most current data to show your location as you move. [21] The "continuous collection" aspect of big data has important consequences both for the technology needed to store the data and the ways that the data can be analyzed, implicating consi-derations of scale, timeliness, privacy, completeness, and accuracy.[22] In fact, velocity may be "perhaps the most challenging

---

[16] *Id.* at 4.

[17] *Id.*; *see also* F.T.C., BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION? UNDERSTANDING THE ISSUES 3-4 (2016) [hereinafter FTC BIG DATA REPORT] https://www.ftc.gov/system/files /documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106 big-data-rpt.pdf; Alex Rosenblat et al., *Networked Employment Discrimination* (Data & Society Research Institute, Working Paper Oct. 8, 2014), https://datasociety.net/pubs/fow/Employ mentDiscrimination.pdf; Matthew T. Bodie et al., *The Law and Policy of People Analytics*

2 (St. Louis Univ. Sch. Of Law Legal Studies Res. Paper Series, Paper No. 2016-6), available at http://ssrn.com/abstract= 2769980.

[18] EOP MAY 2014 REPORT, *supra* note 1, at 4.

[19] *Id.* at 5.

[20] Sarah Guilfoyle et al., *Social Media, Big Data, and Employment Decisions: Mo' Data, Mo' Problems?*, in SOCIAL MEDIA IN EMPLOYEE SELECTION AND RECRUITMENT 127, 131(Richard N. Landers and Gordon B. Schmidt eds., 2016).

[21] EOP MAY 2014 REPORT, *supra* note 1, at 5.

[22] Guilfoyle et al., *supra* note 20, at 131.

component of big data, the ability to manage, and make sense out of information that is continually being collected."[23]

2. Predictive Analytics

The almost incomprehensible volume of data that is rapidly generated from a variety of sources on a continuous basis can be harnessed by a process called predictive analytics.[24] In essence, big data has the capacity to reveal patterns and relationships that would not be visible in a smaller sample size. Predictive analytics uses a method known as data mining to identify trends, patterns, or relationships among data, which can in turn be used to develop a model for predicting behavior based on probabilities.[25] Data brokers compile information from multiple digital and analog sources, unbeknownst to their subjects.[26] Data mining algorithms can be trained to find patterns through the process of "supervised learning," in

which an example of the pattern to be recognized is introduced to the algorithm, or "unsupervised learning," in which the algorithm attempts to identify related pieces of data.[27] Data mining "automates the process of discovering useful patterns, revealing regularities upon which subsequent decision making can rely."[28] "The accumulated set of discovered relationships is commonly called a 'model,' and these models can be employed to automate the process of classifying entities or activities of interest, estimating the value of unobserved variables, or predicting future outcomes."[29] While data mining can identify relationships between seemingly disparate pieces of information, these relationships do not always establish causality.[30]

Predictive analytics has many public and commercial applications. Federal, state, and local governments collect data that can then be used to help improve public services[31] or make the public aware

---

[23] *Id.*

[24] *See* Robert Sprague, *Welcome to the Machine: Privacy and Workplace Implications of Predictive Analytics*, 21 RICH. J. L. & TECH. 13, 1 (2015).

[25] *Id.* at 1, 4.

[26] FTC BIG DATA REPORT, *supra* note 17, at 13 (noting Spokeo "assembled personal information from hundreds of online and offline data sources, including social networks, and merged that data to create detailed personal profiles, including name, address, age range, hobbies, ethnicity, and religion").

[27] EOP BIG DATA AND PRIVACY, *supra* note 8, at 24.

[28] Solon Barocas and Andrew B. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 677 (2016).

[29] *Id.*

[30] *See* EOP BIG DATA AND PRIVACY, *supra* note 8, at 25.

[31] *See, e.g.,* Phil Simon, *Potholes and Big Data: Crowdsourcing Our Way to Better Government*, WIRED (Mar. 2014), https://www.wired.com/insights/2014/03/potholes-big-data-crowdsourcing-way-better-government.

of potential hazards such as consumer product recalls or workplace accidents. [32] However, the vast majority of big data is ultimately used for commercial purposes. After being collected from various sources, it is sold by data brokers to companies for marketing and other purposes. [33] "Data brokers gather not only consumers' spending and debt histories, but also much more intimate details of consumers' financial, social, and personal lives. They track where consumers shop, what they shop for, how they pay for purchases, and much more." [34] That information is then often used to predict consumer behavior, segment consumers into categories for marketing purposes, and generate consumer "scores" that can then be sold to companies to help them determine how to market to or provide services to individual consumers. [35] These scores can determine what ads a customer is shown and even what price they pay. [36] The scores also have important implications for credit eligibility and consumer credit scores.[37]

II. Big Data in the Workplace

A. People Analytics

Big data is now widely used in business, with 62.5% of Fortune 1000 firms reporting at least one application in their business in 2016.[38] While businesses may have first recognized the insights big data can bring to their business and customers, they are also increasingly likely to utilize big data in the area of Human Resources through a subset of predictive analytics known as "people analytics." [39] A 2015 study of 279 members of the Society of Human Resources Management (SHRM)

---

[32] *See, e.g.,* Sandy Smith, *Big Data: OSHA is Poised to Create Massive Data Set of Workplace Injuries and Illnesses*, EHS TODAY (May 11, 2016), http://ehstoday.com/osha/big-data-osha-poised-create-massive-data-set-workplace-injuries-and-illnesses.

[33] *See, e.g.*, F.T.C., DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY 11 (2014), [hereinafter FTC DATA BROKERS], available at https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

[34] Amy J. Schmitz, *Secret Consumer Scores and Segmentations: Separating "Haves" from*

"Have-Nots," 2014 MICH. L. REV. 1411, 1412 (2014).

[35] *Id.* at 1413-1414.

[36] *See generally* Ariel Ezrachi and Maurice E. Stucke, *The Rise of Behavioural Discrimination*, 37 EUR. COMPETITION L. REV. 484 (2016).

[37] *See generally* Schmitz, *supra* note 34.

[38] NEWVANTAGE PARTNERS LLC, *Big Data Executive Survey 2016: An Update on the Adoption of Big Data in the Fortune 1000* 4, (2016), available at http://newvantage.com/wp-content/uploads/2016/01/Big-Data-Executive-Survey-2016-Findings-FINAL.pdf.

[39] Bodie et al., *supra* note 17, at 1-2.

found that while 32% of Human Resources professionals reported that their organization already uses big data to support Human Resources, 82% of organizations planned to either begin or increase their use of big data in Human Resources in the next three years.[40]

Those surveyed by SHRM in 2015 anticipated a number of applications of big data to Human Resources challenges, including evaluating the effectiveness of recruitment campaigns; gauging employee morale and expected retention; making promotion decisions; identifying internal mentors; and locating information across internal applications.[41]

Big data in the Human Resources context means "the combination of nontraditional and traditional employment data with technology-enabled analytics to create processes for identifying, recruiting, segmenting and scoring job candidates and employees."[42]

Nontraditional employment data comes from sources other than the typical personnel data setting, such as "operations and financial data systems maintained by the employer, public records, social media activity logs, sensors, geographic systems, internet browsing history, consumer data-tracking systems, mobile devices, and communications metadata systems."[43] Employers may collect this information internally or may purchase it through data brokers.[44] When combined with traditional employment data like performance reviews, employee longevity, attendance, absenteeism, and salaries, patterns emerge which can then be used to create predictive profiles.[45] Employers can then use these profiles to predict outcomes for job candidates and employees with similar profiles and can deploy these insights in nearly every aspect of the human resources life cycle, including recruitment, hiring, promotion, compensation, and benefit management.[46]

---

[40] THE ECONOMIST INTELLIGENCE UNIT, *Use of Workforce Analytics for Competitive Advantage*, SHRM FOUNDATION 12 (2016), [hereinafter SHRM FOUNDATION], available at http://whitepaper-admin.eiu.com/future hrtrends/wp-content/uploads/sites/2/2016/06/Use-of-Workforce-Analytics-for-Competitive-Advantage.pdf.

[41] *Id.*

[42] *Big Data in the Workplace: Examining Implications for Equal Employment Opportunity Law*, Meeting of the Equal Employment Opportunity Commission (Oct. 13, 2016) [hereinafter *EEOC Big Data Meeting*] (statement of Dr. Kelly Trindel, Chief Analyst, Office of Research, Information, and Planning, EEOC), available at https://www.eeoc.gov/meetings/10-13-16/trindel.cfm.

[43] *Id.*

[44] *Id.*

[45] *Id.*

[46] *Id.*

B. Risks and Opportunities

1. Uncharted Territory

As will be discussed throughout this section, big data and people analytics provide companies with opportunities to achieve business objectives, increase employee wellness, and boost morale. However, companies that seek to channel the potential of big data for lawful purposes must navigate through largely uncharted territory. Existing statutory schemes do not seamlessly apply to big data and people analytics issues. This leaves employers and attorneys who advise them in a regulatory vacuum with little guidance on compliance matters.

Several government agencies and the executive branch have begun to examine the policy questions generated by the increasing use of big data in the employment context. In 2014 and 2016, then-President Barack Obama commissioned a study of these policy questions that gave rise to a series of reports and recommendations. [47] The Federal Trade Commission has also examined big data in consumer credit and hiring, with related privacy considerations.[48]

However, the Equal Employment Opportunity Commission (EEOC) has, perhaps predictably, been the most active governmental entity in probing the legal implications of big data use in the workplace. The first official inquiry by the EEOC into big data issues was at a March 2014 EEOC public meeting entitled "Social Media in the Workplace: Examining Implications for Equal Opportunity Employment Law."[49] That meeting was organized by EEOC Commissioner Victoria Lipnic, [50] who has since been named the agency's Acting Chair. Lipnic identifies the 2014 meeting as the genesis of the EEOC's interest in big data uses, saying that "[s]ince then, all the agency's offices have been educating themselves on the potential bias issues raised by employers' reliance on algorithms and other online tools" and that "the agency is 'very much trying to

---

[47] *See, e.g.*, EOP MAY 2014 REPORT, *supra* note 1; EOP BIG DATA AND PRIVACY, *supra* note 8; EXEC. OFFICE OF THE PRESIDENT, BIG DATA: A REPORT ON ALGORITHMIC SYSTEMS, OPPORTUNITY, AND CIVIL RIGHTS (2016), [hereinafter BIG DATA AND CIVIL RIGHTS], available at https://obamawhitehouse.archives.gov/sit es/default/files/microsites/ostp/2016_05 04_data_discrimination.pdf.

[48] *See, e.g.*, FTC BIG DATA REPORT, *supra* note 17; FTC DATA BROKERS, *supra* note 33.

[49] Kevin McGowan, *When is Big Data Bad Data? When It Causes Bias*, BLOOMBERG BNA (July 29, 2016), https://www.bna.com/big-data-bad-n73014445584.

[50] *Social Media in the Workplace: Examining Implications for Equal Employment Opportunity Law*, Meeting of the Equal Employment Opportunity Commission (Mar. 12, 2014) (statement of Jacqueline A. Berrien, Chair, EEOC), available at https://www.eeoc.gov/eeoc/meetings/3-12-14/transcript.cfm.

understand what is happening,' what's being created by employers and how it's being done." [51] In keeping with these objectives, the EEOC has devoted substantial time in the past two years to big data issues at several public meetings[52] and, on October 13, 2016, the EEOC held its first public meeting exclusively devoted to the use of big data in the workplace and the implications for equal opportunity employment law. [53] Finally, the

EEOC has targeted the use of big data by employers as a long-term research project in its 2016-2019 Research and Data Plan[54] and has identified "the increasing use of data-driven selection devices" in recruitment and hiring as an "area of particular concern" in its Strategic Enforcement Plan for Fiscal Years 2017 through 2021.[55]

These activities represent the initial steps toward examining whether the law needs to be

---

[51] McGowan, *supra* note 49.

[52] *See EEOC at 50: Confronting Racial and Ethnic Discrimination in the 21st Century Workplace*, Meeting of the Equal Employment Opportunity Commission (Apr. 15, 2015) (statement of Dr. Kathleen A. Lundquist, President and CEO, APT Metrics, Inc.), available at https://www.eeoc.gov/eeoc/meetings/4-15-15/lundquist.cfm; *EEOC at 50 July Meeting*, *supra* note 10; *Promoting Diverse and Inclusive Workspaces in the Tech Sector*, Meeting of the Equal Employment Opportunity Commission (May 18, 2016), available at https://www.eeoc.gov/eeoc/meetings/5-18-16/transcript.cfm.

[53] *See EEOC Big Data Meeting*, *supra* note 42.

[54] *Equal Employment Opportunity Commission Research and Data Plan for 2016-2019,* available at https://www.eeoc.gov/eeoc/plan/research_data_plan.cfm. The EEOC offers the following details regarding planning for future big data research:

> The Commission shall consider, as resources allow, the following long-term research projects:…**Research screening devices, tests, and other practices to identify barriers to opportunity across employers and industries as well as promising selection practices**

that rely on job-related criteria:…Study emerging online screening and selection devices, including internet-based assessments that rely on big data analytics and new technology in order to assess the likelihood of employment discrimination caused by these devices. Study validation evidence to determine the likely job relatedness of the instruments used. In addition, develop a centralized bank of information concerning policies, practices, employment inquiries, and employment tests that have raised concerns of discrimination in EEOC investigations. This research will assist the Commission in 1) focusing its enforcement efforts on selection devices and practices that serve as significant barriers to opportunity and 2) providing technical assistance to employers to highlight promising practices and ensure that selection practices focus on job related factors.

(Emphasis original).

[55] *U.S. Equal Employment Opportunity Commission Strategic Enforcement Plan Fiscal Years 2017-2021*, available at https://www.eeoc.gov/eeoc/plan/sep-2017.cfm.

updated to address new issues occasioned by the increasing use of big data in the workplace. Until the law is revised to keep pace with technology, however, counsel should contemplate how existing legal frameworks apply to big data.

2. Compliance Issues to Consider

a. Big Data and the Hiring Process—Complying with the Fair Credit Reporting Act

SCENARIO: *ABC Corp. owns and operates several fast food franchises. It is seeking an Accounts Receivable Clerk for its Accounting Department. In addition to the standard credit check that ABC runs through a credit bureau, ABC intends to review profiles of each applicant prepared by an entity called Worker Profile Co. Worker Profile Co. develops these profiles using information compiled by a data broker. While the credit check is disclosed to the applicants, the profile review is not, as ABC believes that only credit checks*

performed by a credit bureau must be disclosed under federal and state law. Joe Smith applies for the position and passes the credit check. Unbeknownst to Joe, however, ABC decides not to hire him based on information in the profile from Worker Profile Co.

Beginning in the 1990s, companies began to increasingly rely upon online application systems to hire for job openings.[56] Resume database websites enabled candidates to apply for greater numbers of jobs, while allowing employers access to larger candidate pools. [57] Companies turned to analytical tools to score the larger pool of candidates and identify the most qualified individuals. [58] A new frontier for data analytics includes candidate scores in which a profile is built of the ideal candidate for a position and the prospective candidates are scored against the profile. These new scores are built using data about characteristics and behaviors that are well outside traditional factors like education and work experience.[59]

---

[56] *See, e.g.*, BIG DATA AND CIVIL RIGHTS, *supra* note 47, at 13.

[57] *See id.*

[58] *See id.*

[59] *See* Pauline T. Kim and Erika Hanson, *People Analytics and the Regulation of Information under the Fair Credit Reporting Act* (St. Louis Univ. Sch. Of Law Legal Studies Res. Paper Series, Paper No. 16-07-

05) (forthcoming 61 ST. LOUIS U. L.J. 2017), available at https://papers.ssrn.com/ sol3/papers.cfm?abstract_id=2809910, citing Michael Fertik, *Your Future Employer Is Watching You Online. You Should Be, Too*, HARV. BUS. REV. (Apr. 3, 2012), https://hbr.org/2012/04/your-future-employer-is-watchi; *see also* Jeanne Meister, *2014: The Year Social HR Matters*,

Employee recruitment is one of the most common ways in which employers use big data, whether to determine which passive candidates to market to, which candidates to interview, or, ultimately, which applicants to hire. Software applications can find and pair candidates to employers' job postings by looking for the presence of certain words in the candidates' applications, resumes, and social media accounts. [60] LinkedIn's "Talent Match" feature is one such example. These services can use big data to recommend candidates to employers through the use of a training profile or through more overtly-volunteered employer preferences. [61] Both of these approaches necessarily involve certain assumptions (whether machine-learned or employer expressed) about what makes for a "good" employee.

Big data analysis can also suggest new inputs to employers that may have no obvious bearing on candidate suitability but nevertheless correlate with employee success. For example, the recruiting software company Evolv found the following in an analysis of three million data points from over 30,000 hourly employees: (1) employees who installed newer web browsers (such as Mozilla's Firefox or Google's Chrome) perform better and stay longer than the counterparts who used the default browser that came with their computer (usually Internet Explorer for Windows and Safari for Mac); (2) employees who belonged to one or two social networks stayed in their positions longer than those who belong to four or more networks; and (3) employees who live 0-5 miles from their workplace remain in their positions 20% longer than those who live further away. [62] The patterns only become visible because of the large amount of data, although employers and data scientists can only guess why these seemingly irrelevant correlations exist.

The use of algorithms in recruitment can generate both positive and negative results. On the positive side, their use can

FORBES (Jan. 6, 2014), http://www.forbes.com/sites/jeannemeister/2014/01/06/2014-the-year-social-hrmatters/#4b005ef362dc; Don Peck, *They're Watching You at Work*, THE ATLANTIC (Dec. 2013), available at http://www.theatlantic.com/magazine/archive/2013/12/theyre-watching-you-at-work/354681/.

[60] Erin Connell and Mark Thompson, *Big Data, Big Problems: The Liability Pitfall Lurking Beneath the Shiny Surface of "People Analytics"*, ORRICK EMPLOYMENT LAW AND LITIGATION BLOG (Nov. 3, 2015), http://blogs.orrick.com/employment/2015/11/03/big-data-big-problems-the-liability-pitfall-lurking-beneath-the-shiny-surface-of-people-analytics.

[61] Barocas and Selbst, *supra* note 28, at 683.

[62] Alex Rosenblat et al., *Data & Civil Rights: Employment Primer* 2 (Data & Society Research Institute, Produced for Data & Civil Rights Conference Oct. 30, 2014), http://www.datacivilrights.org/pubs/2014-1030/Employment.pdf.

reduce the potential for "like me" bias, in which individuals are most inclined to hire candidates most like themselves. [63]   However, erroneous information obtained from data brokers can deprive an otherwise qualified individual of employment opportunities. [64] Factors such as commuting distance to work, length of time since last job, credit worthiness, and criminal history can compromise the validity of the algorithm's result if they are not aligned with an applicant's qualifications for the job.[65]

The Federal Trade Commission ("FTC") has sought to address this issue through its enforcement of the Fair Credit Reporting Act ("FCRA"). The FCRA applies to entities, called "consumer reporting agencies," that compile and sell consumer reports that companies use to make decisions about a consumer's eligibility for credit, employment, insurance, and housing. [66] The

FCRA's purpose is to ensure that consumer reporting agencies develop reasonable procedures for meeting the needs of companies for access to information "in a manner which is fair and equitable to the consumer, with regard to the confidentiality, accuracy, relevancy, and proper utilization of such information . . . ." [67] Consumer reporting agencies must implement reasonable procedures to ensure accuracy of the information contained within their reports and to provide consumers with access to information, as well as the opportunity to correct any inaccuracies. [68] To this end, the FCRA contains certain notice and consent procedures that companies must adhere to whenever a "consumer report" is used to determine an individual's eligibility for employment and other purposes enumerated in the statute. [69] For example, employers who request a consumer report

---

[63] *See* BIG DATA AND CIVIL RIGHTS, *supra* note 47, at 14.

[64] *See* Kim and Hanson, *supra* note 59, at 2 (noting that "when algorithms rely on error-ridden personal data, they may make inaccurate predictions that arbitrarily reduce individuals' employment opportunities").

[65] *See* BIG DATA AND CIVIL RIGHTS, *supra* note 47, at 14.

[66] *See generally* 15 U.S.C. § 1681 *et seq.*; *see also* FTC BIG DATA REPORT, *supra* note 17, at ii; Mikella Hurley and Julius Adebayo, *Credit Scoring in the Era of Big Data*, 18 YALE J. L. & TECH. 147, 186 (2016) (applicability of the FCRA turns not upon the origin and nature

of the information, but rather the purposes for which the information is collected and the actual or likely end uses of the information).

[67] 15 U.S.C. §1681(b).

[68] FTC BIG DATA REPORT, *supra* note 17, at 13.

[69] *See* 15 U.S.C. §1681 *et seq.* The FCRA defines "consumer report" to mean "any written, oral, or other communication of any information by a consumer reporting agency bearing on a consumer's credit worthiness, credit standing, credit capacity, character, general reputation, personal characteristics, or mode of living . . ." that is used to determine eligibility for

from a consumer reporting agency without first obtaining the individual's consent are in violation of the FCRA.[70] Further, an individual must be issued a "pre-adverse action notice" by any user that intends to take adverse action[71] against a consumer based on information contained within a consumer report.[72] An employer that decides not to hire an individual or to withdraw a conditional offer of employment based on information in a consumer report must provide the consumer with notice of the adverse action and provide the consumer with an opportunity to correct any erroneous information in the report.[73] A covered entity that fails to comply with the FCRA's requirements is subject to civil liability and administrative penalties.[74]

Traditionally, consumer reporting agencies under the FCRA include credit bureaus, employment background screening companies, and other companies

that provide employers with consumer reports that are used to determine a consumer's eligibility for employment.[75] As the use of data brokers to perform predictive analytics expands, so too might the scope of the term "consumer reporting agency."

The FTC has entered into consent decrees with data brokers that advertise their services for employment screening purposes.[76] In one such case, the Federal Trade Commission contended that Spokeo, Inc., a data broker that compiles and sells detailed profiles of consumers, violated the FCRA by marketing the profiles to employers and recruiters for applicant screening purposes. The FTC alleged that Spokeo failed to adhere to key requirements of the FCRA, such as maintaining reasonable procedures to verify the permissible uses of its reports, to ensure the accuracy of its reports, and to provide statutorily required notices to users of its reports.[77] Spokeo did not admit liability;

---

employment and other purposes under the statute.

[70] 15 U.S.C. §1681(d).

[71] The statute defines "adverse action" to include "a denial of employment or any other decision for employment purposes that adversely affects any current or prospective employee." 15 U.S.C. § 1681a(k)(1)(B)(ii).

[72] *See* 15 U.S.C. §1681(m).

[73] *See id.*

[74] *See* 15 U.S.C. § 1681(n)-(s).

[75] *See generally* FTC Big Data Report, *supra* note 17.

[76] *See id.* at ii; *see also* Edward Wyatt, *U.S. Penalizes Online Company in Sale of Personal Data*, N.Y. Times (Jun. 12, 2012), available at http://www.nytimes.com/2012/06/13/te chnology/ftc-levies-first-fine-over-internet-data.html.

[77] Press Release, Fed. Trade Comm'n, *Spokeo to Pay $800,000 to Settle FTC Charges: Company Allegedly Marketed Information to Employers and Recruiters in Violation of FCRA* (Jun.12, 2012), available at https://www.ftc.gov/news-events/press -releases/2012/06/spokeo-pay-800000-

however, it entered into a consent decree that required it to pay an $800,000 civil penalty and enjoined it from further violations of the FCRA. [78] This and similar enforcement actions demonstrate that the FTC takes a broad view of "consumer reporting agency" under the FCRA. Data brokers that compile non-traditional inform-ation, such as social media activity information, may be a consumer reporting agency subject to the FCRA. [79] When an employer purchases predictive analytics services from a data broker for use in making employment eligibility decisions, the employer arguably must comply with the consent and pre-adverse action notice requirements of the FCRA. [80] The FTC has called on data brokers to provide consumers with access to their information through online tools.[81]

In the scenario that introduced this section, Worker Profile Co. could potentially be deemed a consumer reporting agency. Consequently, ABC could potentially be liable for violations of the FCRA if it is using information provided by Worker Profile Co. without adhering to the FCRA's disclosure, authorization, and pre-adverse action notice requirements.

b. Big Data and Disparate Impact Concerns under Title VII of the Civil Rights Act

SCENARIO: Emilia submits an online application for a job with Shipping Brothers, Inc., a logistics company based in downtown Pleasantville, USA. Emilia, who is Hispanic, lives approximately 20 miles away from Shipping Brothers in a suburb of Pleasantville called Woodland Hills. Woodland Hills is

---

settle-ftc-charges-company-allegedly-marketed.

[78] *See* Consent Decree and Order for Civil Penalties, Injunction and Other Relief, United States v. Spokeo, Inc., 2:12-cv-05001-MMM-SH (C.D. Cal. Jun. 9, 2012), available at https://www.ftc.gov/sites/default/files/documents/cases/2012/06/120612spokeoorder.pdf.

[79] *See* FTC BIG DATA REPORT, *supra* note 17, at 13-14.

[80] *See id.* at 15. Note, however, that if the company is using the data to make decisions about its general policies, and not a consumer's eligibility for employment, then the FCRA likely does not apply. *See id.* at 17. Further, information collected on the

activities of a household, neighborhood, or device might not be subject to the FCRA because the information is not collected about "an identifiable person." *See* Hurley and Adebayo, *supra* note 66, at 185; 15 U.S.C. §1681a(c)(d)(1).

[81] *See* FEDERAL TRADE COMMISSION, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: RECOMMENDATIONS FOR BUSINESSES AND POLICYMAKERS 68-70 (2012), available at https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.

primarily populated by Hispanic individuals, like many of Pleasantville's outer suburbs. Shipping Brothers has hired an HR analytics firm, BigDataCo, to assist them with culling the thousands of applications it receives. One of Shipping Brothers' major goals in their current recruitment cycle is to improve employee longevity and retention. BigDataCo has crunched the numbers on employee longevity among Shipping Brothers' existing employees and found that employees who live less than 15 miles away from the company stay in their jobs 20% longer than those who live more than 15 miles away. On this basis, BigDataCo filters out all resumes that include zip codes further than 15 miles from Shipping Brothers, including Emilia's resume and most other Hispanic applicants.

Employers not only have many sources of information about potential candidates but can segment that information in new ways. Employers using big data for their Human Resources needs

should be aware of the potential for discrimination even when using seemingly neutral inputs. They should also consider how internal data bias from third parties can impact employee recruitment, candidate evaluation, and hiring. Advocates of predictive analytics in the hiring process argue that algorithmic techniques eliminate discriminatory bias from the decision-making process. [82] Commentators Barocas and Selbst have argued, however, that because predictive analytics requires "generating a model in which there are winners and losers," it can potentially result in "disproportionately adverse outcomes concentrated within historically disadvantaged groups in ways that look a lot like discrimination."[83]

Some commentators have raised the concern that individuals who are less "data-fied" than other members of society - whether due to poverty, geography, or lifestyle - will be systemically omitted from data sets used to design models.[84] Barocas and Selbst argue that even where data miners are careful to

---

[82] *See generally* Barocas and Selbst, *supra* note 28.

[83] *Id.* at 673.

[84] *See id.* at 684-685 (noting historically disadvantaged groups can be omitted from data miners' data collection an assessment activities because "they are less involved in the formal economy and its data-generating activities, have unequal access to and relatively less fluency in the technology

necessary to engage online, or are less profitable customers or important constituents and therefore less interesting targets of observation"), citing Jonas Lerman, *Big Data and Its Exclusions*, 66 STAN. L. REV. ONLINE 55, 57 (2013); Kate Crawford, *Big Data*: *Why the rise of machines isn't all it's cracked up to be*, FOREIGN POL'Y (May 10, 2013), available at http://foreignpolicy. com/2013/05/10/think-again-big-data/.

address statistical biases, "they can still effect discriminatory results with models that, quite unintentionally, pick out proxy variables for protected classes." [85] Commentators have argued that implicit discrimination can occur when a scoring mechanism includes proxies for race or other protected characteristics as part of its protocol.[86]

The value of data mining ultimately turns upon the quality of the data from which it attempts to draw conclusions. [87] Some commentators have raised concerns that an algorithm can yield discriminatory results depending upon the nature of the data inputted into the algorithm.[88] Data inputs concerns include poorly selected data; incomplete, incorrect, or outdated data; selection bias, where the set of data inputs are not representative of a population and results in a conclusion that could favor certain groups over others; and unintentional promotion of

historical biases.[89] It has also been noted that an algorithm designed to identify candidates that will fit within the existing culture of the company may inadvertently perpetuate past hiring to the exclusion of historically disadvantaged groups.[90] In a May 2016 report to former President Barack Obama, the Executive Office of the President's Big Data Working Group described the issue as follows:

> In a workplace populated primarily by young white men, for example, an algorithmic system designed primarily to hire for culture fit (without taking into account other hiring goals, such as diversity of experience and perspective) might disproportionately recommend hiring more white men because they

---

[85] Barocas and Selbst, *supra* note 28, at 675.
[86] Tal Z. Zarsky, *Understanding Discrimination in the Scored Society*, 89 WASH. L. REV. 1375, 1389 (2014), citing Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 4 (2014); *see also* Barocas and Selbst, *supra* note 28, at 712-714.
[87] *See id.* at 687 (postulating that if data mining incorporates prejudicial or biased behavior of prior decision makers or fails to serve as a good sample of a protected group, it will reach flawed conclusions that could

serve as a discriminatory basis for future decision making).
[88] *See* BIG DATA AND CIVIL RIGHTS, *supra* note 47, at 7; Barocas and Selbst, *supra* note 28, at 684-687 (identifying discrimination concerns arising from incorrect, impartial, or non-representative data).
[89] *See* BIG DATA AND CIVIL RIGHTS, *supra* note 47, at 7-8.
[90] *See id.* at 8 (identifying "unintentional perpetuation and promotion of historical biases" as a potential discriminatory output that can result from the use of big data algorithms in hiring).

score best on fitting in with the culture.[91]

In addition, flaws in the algorithmic system design and interpretation of related data can raise concerns about potential disparate impact.[92]

Other commentators have raised concerns that data mining activities can expose characteristics that people may view as personal and confidential. [93] For example, one study combining data on Facebook "Likes" with other limited information about the subjects was able to predict a male user's sexual orientation 88% of the time; a user's ethnic origin 95% of the time; whether a user was Christian or Muslim 82% of the time; whether a user was Democrat or Republican 85% of the time; and whether the subject used alcohol, drugs, or cigarettes between 65% and 75% of the time. [94] Data mining can uncover sensitive information that should not be considered in the hiring process.

Another example of the utilization of big data in candidate recruitment is the realm of online advertising. There have been several studies in this area, mostly centering on Google's advertising technology because of its relative market dominance.[95] For example, a 2013 study found that Google searches involving black-identifying names (e.g., Darnell) were more likely to display ads with the word "arrest" in them than searches with white-identifying names (e.g., Geoffrey). [96] The opacity of Google's search function prevented the researcher from determining exactly why this result occurred because choices about ad display (like other big data calculi) are determined by complex algorithms comprised of multiple internal decision processes. [97] A second study, conducted in 2014, found that when Google perceived a searcher's gender to be female, the searcher was shown fewer ads for high paying jobs than when Google perceived the searcher's gender to

---

[91] *See id.*

[92] *See id.* at 8-10 (identifying poorly designed matching systems; personalization and recommendation services that narrow rather than expand user options; decision-making systems that equate correlation with causation; and data sets that lack information about or over-represent certain populations as design flaws that can be imbedded in algorithmic systems).

[93] FTC BIG DATA REPORT, *supra* note 17, at 10.

[94] *Id.*, citing Michal Kosinski, et al., *Private Traits and Attributes Are Predictable from*

Digital Records of Human Behavior, 110 PROCEEDINGS OF THE NAT'L ACAD. OF SCI. 5802, 5803-5804 (2013).

[95] *See, e.g.,* Latanya Sweeney, *Discrimination in Online Ad Delivery*, COMMC'NS OF THE ACM (May 2013); Amit Datta et al., *Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination*, PROCEEDINGS ON PRIVACY ENHANCING TECHNOLOGIES (Apr. 2015).

[96] Sweeney, *supra* note 95, at 11.

[97] EOP MAY 2014 REPORT, *supra* note 1 at 7.

be male.[98] Once again, because of the opacity of the "ad ecosystem, which includes Google, advertisers, websites, and users," researchers could not determine why these findings occurred.[99]

Title VII of the Civil Rights Act of 1964 prohibits discrimination against individuals based on certain characteristics protected by law, such as race, gender, or religion.[100] Under Title VII, there are two theories of discrimination: (1) disparate treatment and (2) disparate impact. Disparate treatment involves cases of intentional discrimination in which the employer treats an employee differently than similarly situated employees because of his protected characteristic. Disparate impact involves cases where policies or practices that are facially neutral have a disproportionate adverse effect on protected classes.[101] Discrimination in the disparate treatment context focuses on the decision maker's state of mind to answer the question of whether the decision maker intended to discriminate against an individual because of his or her membership in a protected class.[102]

Discrimination claims arising out of the use of predictive analytics to make hiring decisions lend themselves more to the disparate impact theory of liability because the algorithms themselves will rarely be intentionally designed to assess a potential applicant's protected characteristics. For the reasons discussed below, however, the disparate impact theory might also prove to be ill-suited to a claim based on the use of predictive analytics to make employment decisions.

Under Title VII, a plaintiff in a disparate impact case need not establish that the employer acted with a discriminatory motive or intent. [103] Rather, the employee must establish that the facially neutral policy or practice had a disparate impact with respect to a protected class.[104] If the employer demonstrates that the challenged policy or practice is job-related and consistent with business necessity, then the plaintiff must establish a viable "alternative employment practice" that would have a less

---

[98] Datta et al., *supra* note 95, at 1.

[99] *Id.*

[100] *See* 42 U.S.C § 2000e-2(k).

[101] *See* Griggs v. Duke Power Co., 401 U.S. 424, 430 (1971).

[102] *See* Zarsky, *supra* note 86, at 1382 (noting that "discussing discrimination in the context of the scored society challenges existing thought paradigms").

[103] Notably, a disparate impact claim can succeed even where the employer did not intend to discriminate. *See* Jones v. City of Boston, 752 F.3d 38, 46 (1st Cir. 2014), citing Boston Chapter, N.A.A.C.P., Inc. v. Beecher, 504 F.2d 1017, 1021 (1st Cir. 1974).

[104] *See* 42 U.S.C. § 2000e-2(k)(1)(A).

discriminatory impact on the protected class.[105]

Plaintiffs asserting discrimination claims arising out of the use of predictive analytics will likely face obstacles to establishing the claim where the algorithm uses facially non-discriminatory factors that correlate with strong job performance. When assessing whether the challenged practice is "job-related and consistent with business necessity," courts assess whether the challenged practice is predictive of future job performance.[106] However, there is generally no requirement that the challenged practice be essential or indispensable to the employer's business for it to be job-related and consistent with business necessity. [107] Assuming that the employer's business justification passes muster, then the plaintiff must identify an alternative practice that serves the business purpose without producing the same adverse impact on protected groups.[108] In the big data context, to prevail under a disparate impact theory of discrimination a plaintiff must show the algorithm used to make an employment decision adversely impacts a protected group or, if the employer succeeds in establishing legitimate business reasons for using the algorithm, by demonstrating there exists an alternative that is equally efficient at serving the employer's legitimate business needs.[109] Employers who use big data for lawful means can argue that the intent of using predictive analytics in hiring decisions is tied to legitimate business purposes - improving productivity, retention, and profitability.

Further, given the opaque nature of predictive analytics, plaintiffs will likely face challenges in proffering statistical evidence of discrimination. In order to establish a *prima facie* case of employment discrimination based on disparate impact theory, the plaintiff must: "(1) identify a specific employment practice that is being challenged and (2) establish, through statistical means, that the identified employment practice 'caused the exclusion of

---

[105] *See id.* While the statute does not define the threshold showing required to establish disparate impact, the EEOC has developed the so-called "four-fifths rule." In the Uniform Guidelines on Selection Procedures, the EEOC states that "[a] selection rate that is less than four-fifths . . . of the rate for the group with the highest rate will generally be regarded . . . as evidence of adverse impact." 29 C.F.R. § 1607.4(D)(2015).

[106] *See* Barocas and Selbst, *supra* note 28, at 702 (discussing analysis of "business necessity" under *Griggs* and related jurisprudence).
[107] *See Id.* at 703 (discussing Wards Cove Packing Co. v. Atonio, 490 U.S. 642 (1989)).
[108] *See* Albemarle Paper Co. v. Moody, 422 U.S. 405, 425 (1975).
[109] Marko Mrkonich et al., *The Big Move Toward Big Data in Employment*, THE LITTLER REPORT 3 (2015), available at https://perma.cc/EP56-6NHP.

applicants . . . because of their membership in a protected group.'"[110] A disparate impact case will fail if the plaintiff is unable to provide reliable statistical evidence of discrimination.[111] The Supreme Court has most recently described a prima facie showing of disparate impact as "essentially, a threshold showing of a significant statistical disparity . . . and nothing more."[112] Finally, commentators have also been critical of a sense of "data determinism" in the decision-by-algorithm process, wherein predictive analytics uses correlations to draw inferences and make judgments about individuals based on what they might do, rather than what they have actually done. [113] With respect to hiring, individuals are judged not by what they have done, but because correlations revealed by predictive analytics suggest that they might be unsuitable candidates for employment. [114] This "data determinism" arguably runs counter to the notion of a meritocracy in which individuals are evaluated based on their contributions in the workplace.

c. Reasonable Accommodations in the Hiring Process—The ADA and ADA Amendments Acts

SCENARIO: Peter, who has a disability, has applied for a job at DEF Code Corp., a leading technology company in his area. DEF Code's application portal, which Peter uses to access the company's online application, tracks which type of Internet browser each potential applicant utilizes in navigating to and around its portal. DEF Code tracks which browser an applicant uses because DEF Code has data on existing employees that shows that those who use the browser that comes standard with their PC are less productive than those who download other browsers. Peter uses

---

[110] E.E.O.C. v. Kaplan Higher Educ. Corp., No. 1:10-cv-2882, 2013 WL 322116 (N.D. Ohio Jan. 28, 2013) *aff'd* 748 F.3d 749 (6th Cir. 2014), quoting Watson v. Fort Worth Bank & Tr., 487 U.S. 977, 994 (1988).

[111] *See id.* (granting summary judgment for employer in disparate impact case involving use of credit checks where expert evidence offered to support statistical evidence of discrimination was deemed unreliable).

[112] Ricci v. DeStefano, 557 U.S. 557, 587 (2009); *see also* Fudge v. City of Providence Fire Dep't, 766 F.2d 650, 658 n.8 (1st Cir. 1985) (holding that a prima facie case of disparate impact can be established where "statistical tests sufficiently diminish chance as a likely explanation").

[113] Zarsky, *supra* note 86, at 1408-1409, citing Edith Ramirez, FTC Chairwoman, Keynote Address at the Tech. Policy Inst. Aspen Forum, *The Privacy Challenges of Big Data: A View from the Lifeguard's Chair* 7-8 (August 19, 2013), available at https://www.ftc.gov/sites/default/files/d ocuments/public_statements/privacy-challenges-big-data-view-lifeguard%E2%80%99s-chair/130819bigdataaspen.pdf.

[114] *Id.*

the browser that comes standard with his PC, and, as a result, is excluded from consideration for employment on this basis.

The Americans with Disabilities Act of 1990 as amended by the ADA Amendments Act of 2008 ("2008") prohibits employers from administering tests and selection criteria that exclude individuals with disabilities from jobs that they can actually perform merely because their disability prevents them from taking a test or negatively influences the results of a test that is a prerequisite for the job.[115] The EEOC's Interpretative Guidance on selection criteria and the ADA requires employers to "select and administer tests concerning employment in the most effective manner to ensure . . .[that when it is administered to a disabled individual] the test results accurately reflect the skills, aptitude, or whatever other factor of the applicant or employee that the test purports to measure, rather than reflecting the impaired sensory, manual, or speaking skills of such employee or applicant . . ." Commentators have noted that this guidance, which was written over

twenty years ago, is not readily transferable to the use of big data in employment screening processes.[116] First, the criteria that big data uses to create algorithms, much of which might be unrelated to work activities themselves, is not traditionally regarded as a "test " as that term is currently used in the EEOC guidance.[117] Second, the information relied upon by big data is generated during the course of daily living and collected without an applicant's knowledge. Because of this dynamic, an applicant with a disability will not know to request an accommodation and the employer will not know of the applicant's need for an accommodation.[118]

Consider the scenario involving Peter referenced above. If the reason Peter uses the standard browser is related to the fact that certain features of the browser make it easier to use than other commercially available alternatives for people with his particular disability, DEF Code's use of browser choice as a hiring criteria could involve issues of reasonable accommodation as well as disparate impact considerations.

---

[115] 29 C.F.R. § 1630.11; *see also* 42 U.S.C. § 12112(b)(5)(A)(2012) (requiring reasonable accommodation of qualified applicants with disabilities).

[116] *See* Allan G. King and Marko J. Mrkonich, "*Big Data" and the Risk of Employment*

Discrimination, 68 OKLA. L. REV. 55, 581-582 (2016).

[117] *See id.*

[118] *See id.* at 582.

d. Big Data and Confidential Health Information— Complying with the ADA, GINA, and HIPAA

SCENARIO: Helen Roberts, a Human Resources professional at Toy Fun, Inc., has been tasked with evaluating the administration of Toy Fun's voluntary employee wellness program. One of Toy Fun's goals in adopting a wellness program is to reduce the number of sick days utilized by employees in order to increase profitability and lower insurance premiums. Because Toy Fun is a large employer, it has the capacity to administer the new program in-house, which Helen is excited about because it is less costly than hiring an outside vendor and will allow Helen's staff to design a program that best fits Toy Fun's goals. However, whether the program is administered in-house or by a vendor, Helen is concerned about what (and how) employee

data can permissibly be used to achieve program success while at the same time complying with various federal and state health privacy laws.

Benefits management, a key cost-center for employers, is another area where big data is being deployed. In an effort to reduce healthcare costs and increase productivity by encouraging employees to live healthier lifestyles, many employers are adopting wellness programs.[119] The term "wellness program" generally refers to health promotion and disease prevention programs and activities offered to employees by employers as part of an employer-sponsored group health plan or as a stand-alone program.[120]

Wellness programs frequently ask employees to disclose medical information on health risk assessment questionnaires or to undergo biometric screenings [121]

---

[119] *See* Jay Hancock, *Workplace Wellness Programs Put Employee Privacy at Risk*, CNN (Oct. 2, 2015), http://www.cnn.com/ 2015/09/28/health/workplace-wellness-privacy-risk-exclusive.

[120] EEOC Final Rule on Employer Wellness Programs and Title I of the Americans with Disabilities Act, 29 C.F.R. § 1630.14 (2017); Ifeoma Ajunwa, Kate Crawford, and Joel S. Ford, *Health and Big Data: An Ethical Framework for Health Information Collection by Corporate Wellness Programs*, 44 J.L. MED. & ETHICS 474, 475 (2016) (noting wellness programs began as Employee Assistance Programs designed to assist employees with mental health issues,

substance abuse, and stress, and have evolved to include health risk assessment, weight reduction, and smoking cessation programs, among other things).

[121] The Centers for Disease Control and Prevention's Workplace Health Glossary defines "biometric screening" as "the measurement of physical characteristics such as height, weight, body mass index, blood pressure, blood cholesterol, blood glucose, and aerobic fitness tests that can be taken at the worksite and used as part of a workplace health assessment to benchmark and evaluate changes in employee health status over time." CENTERS FOR DISEASE CONTROL AND PREVENTION, https://www.cdc.

for risk factors.[122] One 2015 study by the Kaiser Family Foundation found that 81% of businesses with 200 or more workers that offered health benefits had a wellness program, with 50% of such employers utilizing health risk assessments and/or biometric screening. [123] A number of apps have been developed to track employee health and fitness data for use by wellness programs. [124] Wearable devices, such as the FitBit, can be used to collect and transmit data regarding an employee's height, weight, heart rate, physical activity, and sleep patterns.[125]

In addition to simply administering the wellness programs, some employee wellness firms and insurers are working with third party data brokers to mine data about the prescription drugs workers use, where and how they shop, and even whether they vote, all in order to predict an individual employee's health needs and recommend treatments.[126] For example, Wal-Mart reportedly engaged a data broker firm, Castlight, to collect and analyze

employee data to determine which workers were at risk for diabetes in order to provide them with targeted preventative health information.[127] According to news reports, it also engaged the third-party firm to review insurance and pharmaceutical claims to identify employees likely to be considering spinal surgery in order to steer them toward less costly treatment options. [128] Castlight has also reportedly determined which employees are likely to become pregnant by scanning insurance claims to find employees who have stopped filling birth control prescriptions or searched for fertility-related topics on Castlight's health app. [129] Employees who fit Castlight's search criteria were provided with information about prenatal care and choosing an obstetrician.[130]

Employee wellness programs have raised several concerns. The stated goal of such practices is a positive one: to improve employee health and, in turn, decrease health

---

gov/workplacehealthpromotion/tools-resources/glossary/glossary.html.

[122] EEOC Final Rule on Employer Wellness Programs and the Genetic Information Nondiscrimination Act, 29 C.F.R. § 1635.8, §1635.11 (2017).

[123] Sharona Hoffman, *Big Data and the Americans with Disabilities Act* 4 (Case Legal Studies Research Paper No. 2016-33), available at https://papers.ssrn.com/sol3/papers2.cfm?abstract_id=2841431.

[124] *See* Hancock, *supra* note 119.

[125] *See id.*

[126] Rachel Emma Silverman, *Bosses Tap Outside Firms to Predict Which Workers Might Get Sick*, WALL ST. J. (February 17, 2016), available at https://www.wsj.com/articles/bosses-harness-big-data-to-predict-which-workers-might-get-sick-1455664940.

[127] *See id*.

[128] *See id.*

[129] *See id.*

[130] *See id.*

care expenses.[131] However, some are concerned about employers using algorithms to analyze electronic health data to determine which individuals might be high cost or less productive workers.[132] Some have argued that wellness programs violate the ADA's prohibition against involuntary medical exams and disability-related inquiries. Others have argued that the provision of medical information concerning family members as part of a health risk assessment violates the Genetic Information Non-Discrimination Act (GINA). Some scholars have commented that the strong financial incentives to participate render these programs essentially involuntary and coerce disclosure of protected health information.[133]

Finally, privacy concerns have arisen about the provision of the information to third-party data brokers[134] and the security of the information maintained by these programs.[135] The EEOC, which is charged with enforcement of the ADA and GINA, sought to address some of these concerns in its Final Rule on Employer Wellness Programs and the Genetic Information Nondiscrimination Act and its Final Rule on Employer Wellness Programs and the Americans with Disabilities Act, both issued on May 17, 2016.

Title I of the Americans with Disabilities Act of 1990 seeks to eliminate disability-based discrimination in the workplace.[136] Among other obligations in the ADA, an employer cannot make disability-related inquiries or require an employee to submit to a medical examination unless they are "job-related and consistent with business necessity."[137] There is an exception to this prohibition for

---

[131] *See id.*

[132] *See* Frank Pasquale and Tara Adams Ragone, *Protecting Health Privacy in an Era of Big Data Processing and Cloud Computing*, 17 Stan. Tech. L. Rev. 595, 636 (2014), citing Sharona Hoffman, *Employing E-Health: The Impact of Electronic Health Records in the Workplace*, 19 Kan. J.L. & Pub. Pol'y, 409, 422 (2010) (discussing possibility of "complex scoring algorithms based on EHRs to determine which individuals are likely to be high-risk and high-cost workers").

[133] Ajunwa et al., *supra* note 120, at 475, citing D.C. Rubenstein, *The Emergence of Mandatory Wellness Programs: Should Your Employer Be the Boss of More Than Your Work?*, 38 Sw. U. L. Rev. 465, 468-469 (2009).

[134] Ajunwa et al., *supra* note 120, at 478.

[135] *See* Ifeoma Ajunwa, *Workplace Wellness Programs Could Be Putting Your Health Data at Risk*, Harv. Bus. Rev. (January 19, 2017), available at https://hbr.org/2017/01/workplace-wellness-programs-could-be-putting-your-health-data-at-risk.

[136] *See* 42 U.S.C. § 12101, *et seq.*

[137] 42 U.S.C. § 12112(d)(4)(A) (a covered entity "shall not require a medical examination and shall not make inquiries of an employee as to whether such employee is an individual with a disability or as to the nature or severity of the disability, unless such examination or inquiry is shown to be job-related and consistent with business necessity"); *see also* EEOC Enforcement Guidance: Disability-Related Inquiries and Medical Examinations of Employees under the Americans with Disabilities Act,

inquiries and examinations made in connection with an "employee health program." [138] A medical inquiry or examination must satisfy several conditions to be acceptable under the employee health program exception. First, the health program must be voluntary. Under the Final Rule, medical inquiries and examinations are "voluntary" when the employer does not "(1) require employees to participate; (2) deny coverage under any group health plan to employees for non-participation; (3) take any adverse action, retaliate against, or coerce employees who choose not to participate."[139]

Further, a health program remains voluntary if the financial penalty for non-participation remains at or below thirty percent of the cost for self-only coverage.[140] In addition to being voluntary, the program must be reasonably designed to promote health or prevent disease. Employees must be provided with notice of the type of medical information that will be obtained and the reason for obtaining it. [141] The information must be maintained in a separate medical file with appropriate safeguards against disclosure of an

employee disability, unless such disclosure is required to provide an accommodation. [142] Generally, however, employee health information should be provided to the employer "in aggregate terms that do not disclose, or are not reasonably likely to disclose, the identity of any employee."[143]

Finally, and, from a privacy standpoint, most importantly, an employer cannot require an employee to agree to the "sale, exchange, sharing, transfer, or other disclosure of medical information (except to the extent permitted by this part to carry out specific activities related to the wellness program), or to waive any confidentiality protections in this part as a condition for participating in a wellness program or for earning any incentive the covered entity offers in connection with such a program."[144]

Title II of the Genetic Information Non-Discrimination Act of 2008 prohibits discrimination on the basis of genetic information in employment. [145] It prohibits employers and other covered entities from using genetic information when making

---

https://www.eeoc.gov/policy/docs/guidance-inquiries.html.

[138] 42 U.S.C. § 12112(d)(4)(B); 29 C.F.R. § 1630.14(d).

[139] EEOC v. Orion Energy Sys., Inc., No. 14-CV-1019, 2016 WL 5107019 (E.D. Wisc. Sept. 19, 2016), citing 81 Fed. Reg. 31,126, 31,133 (May 17, 2016).

[140] *Id.*, citing 81 Fed. Reg. 31,126, 31,134 (May 17, 2016).

[141] *See* 29 C.F.R. § 1630.14(d)(2)(iv).

[142] *See* 29 C.F.R. § 1630.14(d)(4)(i).

[143] 29 C.F.R. § 1630.14(d)(4)(iii).

[144] 29 C.F.R. § 1630.14(d)(iv).

[145] *See* 42 U.S.C. §§ 2000ff *et seq.*; 29 C.F.R. § 1635 (2017).

decisions about employment. [146] The statute and the EEOC's GINA regulations define the term "genetic information" to include information about "manifestation of disease or disorder in the family members of an individual," including blood relatives and other dependents. [147] Often, a health risk assessment administered as part of a wellness program will elicit information about an employee family member's health status, seemingly running afoul of GINA. There is an exception in GINA, however, when employers acquire genetic information as part of a voluntary health or wellness program. [148] The EEOC's Final Rule clarifies that an employer may offer a limited incentive for an employee's spouse to provide information about the spouse's current or past health status as part of a voluntary wellness program without running afoul of GINA, provided that GINA's confidentiality requirements are observed and the any information

obtained is not used to discriminate against an employee. [149] In order to pass muster under the Final Rule, the wellness program must be (1) voluntary; [150]   (2) reasonably designed to promote health or prevent disease; [151]   and   (3) maintain the confidentiality of the employee's genetic information.

The Final Rule also prohibits employers from requiring and employee or spouses to agree to the sale, exchange, transfer, or other distribution of health information in exchange for an inducement or as a condition for participating in the wellness program. [152] Under both Final Rules, employers may offer an employee's adult and minor children the opportunity to participate in a wellness program, but may not offer any inducement in exchange for information about their current or past health status.

In October 2016, the AARP, an association that represents over 38 million people over age 50,[153] filed suit against the EEOC arguing that

---

[146] *Id.*

[147] *See id.*; 42 U.S.C. §§ 2000ff(4), 2000ff-8(b).

[148] *See* 42 U.S.C. § 2000ff-1(b)(2)(A)-(B).

[149] *See id.*; *see also* 29 C.F.R. § 1635.8(b)(2)(i)(A).

[150] As with the Final Rule under the ADA, the inducement cannot exceed 30% of the total cost of self-only coverage in order to remain voluntary.

[151] According to the Final Rule, a wellness program is not reasonably designed to promote health or prevent disease if "it exists merely to shift costs from an employer to employees based on their

health; is used by the employer only to predict its future health costs; or imposes unreasonably intrusive procedures, an overly burdensome amount of time for participation, or significant costs related to medical exams on employees." EEOC's Final Rule on Employer Wellness Programs and the Genetic Information Non-Discrimination Act, https://www.eeoc.gov/laws/regulations/qanda-gina-wellness-final-rule.cfm.

[152] *See id.*

[153] *See* AARP v. EEOC, C/A No. 16-02113-JDB, 2016 WL 7646358, at *6 (D.C. Dec. 29, 2016).

wellness programs violated anti-discrimination laws under ADA and GINA and that the high penalty assessed to employees for non-participation renders them involuntary.[154] The suit sought a preliminary injunction enjoining the implementation of the Final Rule, contending that the Final Rule sanctioned plans that coerce employees into disclosing confidential health information during a health risk assessment, biometric testing, or other health-related inquiries.[155] AARP argued that employees were effectively compelled to reveal disability-related information protected under the ADA and genetic information protected under GINA.[156] Industry groups representing employers have pointed out that health information collected as part of a wellness program is presented to employers in aggregate, de-identified form and that there is no evidence employers are using the information to discriminate against employees on the basis of their disabilities or genetic information.[157] Further, employers engage third-party vendors to administer wellness programs precisely so that they will

not know whether an employee has a medical condition[158]

A federal district court judge denied the request for injunction, finding that the representative plaintiffs had not demonstrated irreparable harm through the payment of higher premiums.[159] Further, while the Final Rule permits disclosure of protected information in some circumstances as part of a wellness program, the statutory provisions of the ADA and GINA prohibit employers from using protected information to discriminate against employees.[160] The Court did note, however, that while the AARP had not met the standard for issuance of a preliminary injunction, the substantive issue of whether the Final Rules were in conformance with the Administrative Procedure Act remained for decision on the full administrative record.[161]

Typically, employees are asked to authorize the third-party vendor to review the protected health information. Some privacy advocates have raised concerns that employee health data disclosed to a third-party vendor administering an employee wellness program can be sold to a third-party data broker and

---

[154] Reed Abelson, *AARP Sues U.S. Over Rules for Wellness Programs*, N.Y. TIMES (Oct. 24, 2016), available at https://www.ny times.com/2016/10/25/business/employ ee-wellness-programs-prompt-aarp-lawsuit.html.
[155] *See id.*

[156] *See AARP*, 2016 WL 7646358, at *1.
[157] *See id.*
[158] *See* Ajunwa et al., *supra* note 120, at 475-476.
[159] *See AARP*, 2016 WL 7646358, at *8-9.
[160] *See id.* at *9.
[161] *See id.* at *12.

incorporated into algorithms used to create profiles of the employee.[162] Even de-identified information can be combined with publicly available data to re-identify wellness program participants.[163] Where a wellness program is offered as part of a group health plan, HIPAA's privacy rule, security rule, and breach notification provisions apply to any employee or dependent information collected as part of the program.[164] However, a question arises as to the confidentiality of the information when the wellness program is administered by a third-party vendor that is not considered to be a covered entity under HIPAA. Further, how secure is the data collected by wellness programs and stored in cloud-based systems? HIPAA mandates data security provisions regarding protected health information in contracts between employers and third party service providers.[165] States have

similar provisions requiring employers to make arrangements with third-party service providers to ensure the security of health-related data.[166] Many more states require employers to protect their employees' personal information.[167] Finally, employers can be required by statute to notify employees in the event of a data breach, even where the breach occurred to the third-party vendor.[168]

e. Workforce Science—Opportunities (and Challenges) for Employers

Big data is also being used to monitor employee productivity in the new field of "workforce science."[169] In workplace science, employers monitor and collect data about employee phone calls, emails, computer use, and other digital and non-digital behavior.[170] Workplace

---

[162] *See* Hancock, *supra* note 119; *see also* Pasquale and Ragone, *supra* note 132, at 630 (providing examples in which data brokers sell information regarding consumer medications and ailments).

[163] *See id.*

[164] *See id.*

[165] 45 C.F.R. 164.5(c).

[166] *See, e.g.*, Cal. Civ. Code § 1789.81.5(c) (West 2017); Mass. Gen. Laws ch. 93H, § 2 (2017) (as implemented by 201 C.M.R. 17.00); Or. Rev. Stat. § 646A.622(2)(d) (2017).

[167] *See, e.g.*, Fla. Stat. Ann. § 501.171(2) (West 2017); Tex. Bus. & Com. Code Ann. § 521.052(a) (West 2017).

[168] *See, e.g.*, 815 Ill. Comp. Stat. 530/5 *et seq.* (2017); Mich. Comp. Laws § 445.72 (2017); N.Y. Bus. Law § 899-aa (McKinney 2017); Ohio Rev. Code Ann. § 1349.19 (West 2017).

[169] Sprague, *supra* note 24, at 31; *see also* Steve Lohr, *Big Data, Trying to Build Better Workers*, N.Y. Times (Apr. 21, 2013), available at http://www.nytimes.com/ 2013/04/21/technology/big-data-trying-to-build-better-workers.html (describing workforce science as "what happens when Big Data meets H.R."); Peck, *supra* note 59.

[170] Lohr, *supra* note 169.

predictive analytics can then be used to analyze this data in an effort to improve productivity, efficiency, and innovation. [171]   Some companies have used big data analytics in an effort to increase employee productivity and morale and to decrease employee turnover.[172] As previously discussed, big data can be used to build profiles of an ideal candidate based on traits that do not appear to have a direct link to better job performance, and companies like Google and I.B.M. are using workforce surveys to identify the traits that are predictive of success in a given position. [173]   The information gleaned from these surveys can then be used to develop tests administered to job applicants. While American corporations have long administered a variety of tests to applicants, the power of big data and predictive analytics has

changed the nature of workforce science. [174]   In addition, commentators have opined that analysis of employee activities can be used to eliminate the subjective nature of performance appraisals, which can reflect implicit or unconscious bias of the decision maker. [175]   Some employers have even used predictive analytics for the desirable purpose of identifying and remedying cases of disparate treatment.[176]

For as much information as employers have about applicants, this data pales in comparison to the data employers have about their current employees. To wit, "Big Data also holds out the promise of, for instance, total supervision in the workplace…Every phone call, email and even mouse-click of an employee can be stored and analyzed to guide management in making decisions." [177]   This infor-

---

[171] Sprague, *supra* note 24, at 33, citing Don Reisinger, *Improving Employee Performance with Data Analysis*, CIO INSIGHT (Aug. 20, 2013), http://www.cio insight.com/it-management/workplace/ slideshows/improving-employee- performance-with-data-analysis and Steve Lohr, *Scientific Management Redux: The Difference Is in the Data*, N.Y. TIMES BLOG (Apr. 21, 2013), https://bits.blogs. nytimes.com/2013/04/21/scientific- management-redux-the-difference-is-in- the-data).

[172] Mrkonich, *supra* note 109, at 12; Steven Pearlstein, *People Analytics: 'Moneyball' for Human Resources*, WASH. POST (Aug. 1, 2014), available at https://www.washingtonpost. com/businesspeople-analytics-moneyball- for-human-resources/2014/08/01/

3a8fb6ac-1749-11e4-9e3b-7f2f110c6265 story.html?utm term=.8ccbff95f432.

[173] Lohr, *supra* note 169.

[174] Lohr, *supra* note 171.

[175] *EEOC Big Data Meeting*, *supra* note 42 (statement of Michal Kosinski, Asst. Prof. of Organizational Behavior at Stanford School of Business), available at https://www. eeoc.gov/eeoc/meetings/10-13-16/ kosinski.cfm.

[176] Ben Waber, *What Data Analytics Says About Gender Inequality in the Workplace*, BLOOMBERG (Jan. 30, 2014), https://www.bloomberg.com/news/articl es/2014-01-30/gender-inequality-in-the- workplace-what-data-analytics-says.

[177] Steven Poole, *Are you ready for the era of Big Data?*, NEW STATESMAN (May 29, 2013), https://perma.cc/2DPN-SBQC.

mation can be used to predict which—and sometimes, when—employees are likely to quit, [178] suffer a workplace accident, [179] or take a medical leave of absence. [180] The data can also provide employers with suggestions for which employees to promote and which to terminate. [181]

One software company, SAS, has developed an employee retention program that analyzes data on employees who have quit, including their "skills, profiles, studies, and friendships." [182] Similarly, the Evolv study mentioned in Section III(B)(2)(a) analyzed preferences for Internet browsers, social network participation, and proximity to work, finding that all of these data points correlated to employee retention and success. [183] Another company, Cataphora, studies intra-company communications by analyzing data samples, whether words or software code, and determining which employees are "thought leaders" (people whose words or work product are copied or cited most frequently) and "networked curators" (those who

perceive valuable content and share it with others). by mapping communications in this way, Cataphora's clients can determine who is most valuable to the company (and who is less so). [184]

Employers are also using tracking and assessing employees' physical locations and in-person communications in addition to their online communications. [185] One data analytics firm, Sociometric Solutions, measured employee in-person collaboration using sensor ID badges which detect conversations and speech patterns via infrared, Bluetooth, and microphone data and monitored physical movement through the use of accelerometers. [186] Because there is some evidence which suggests that eating lunch with one's work colleagues is more productive than eating alone, where (and with whom) an employee eats lunch could be a useful data point to monitor for certain employers. [187]

As one of Sociometric Solutions' founders put it, "[s]ociometrics is all about analyzing the patterns of relationships that connect people. In the workplace, interacting with

[178] Sprague, *supra* note 24, at 33.

[179] *Id.*; *see also* Stephen Baker, *Data Mining Moves to Human Resources*, Bloomberg Businessweek (Mar. 12, 2009), https://www.bloomberg.com/news/articles/2009-03-11/data-mining-moves-to-human-resources; Smith, *supra* note 32.

[180] Aimee Picchi, *The "big data" app that predicts employees' health*, CBS MoneyWatch (Feb. 18, 2016),

http://www.cbsnews.com/news/the-big-data-app-that-predicts-employees-health.

[181] Sprague, *supra* note 24, at 33.

[182] *Id.*

[183] Rosenblat et al., *supra* note 17.

[184] Baker, *supra* note 179.

[185] Sprague, *supra* note 24, at 34.

[186] Waber, *supra* note 176.

[187] Sprague, *supra* note 24, at 34.

the right people in the right way is vital."[188] Quantifying the physical interactions and social engagement of employees has all kinds of applications for employers: they can target top performers for rewards or low performers for remediation (or termination), understand how their workplace's physical design contributes to or detracts from collaboration, and build teams based on employee work styles and communication strengths or weaknesses.

Big data can also help employers better identify and promote internal talent. By combining data sets from across an organization, employers can find qualified candidates who might be in another area of the company, something that can also have the effect of identifying and promoting women and minorities. [189] Employers can also use big data to examine how their internal policies may contribute to or detract from increased diversity. For example, Google, which relies heavily on employee self-nominations for promotions, found that women were less likely to self-nominate, a fact that was depressing the ranks

of women in upper management.[190] Google's solution also lay within its internal data: it discovered that it could successfully "nudge" women to self-nominate simply by sending an email encouraging them to do so.[191]

While the potential for workforce science to improve worker efficiency and satisfaction offers opportunities for organizations to thrive, the employee monitoring raises policy questions. How closely should employees be monitored, both inside and outside of the workplace as well as online?[192] How much information should be disclosed to employees about an employer's monitoring activities? These questions and others raised by emerging technologies present significant challenges to traditional notions of privacy in the employment context.

III.   Conclusion

Big data has the potential to dramatically change the workplace. Employers and their counsel must consider the legal implications of implementing new technology into

---

[188] *Id.*

[189] Bonnie Marcus, *How Big Data is Helping Women Move to the C-Suite*, FORBES (Feb. 23, 2016), https://www.forbes.com/sites/ bonniemarcus/2016/02/23/how-big-data-is-helping-move-women-to-the-c-suite/#3750a5a97799.

[190] Cecilia Kang, *Google data-mines its approach to promoting women*, WASH. POST

(Apr. 2, 2014), https://www. washingtonpost.com/news/the-switch/wp/2014/04/02/google-data-mines-its-women-problem/?utm_term=.ef9a34bfe94c.

[191] *Id.*

[192] *See* SHRM FOUNDATION, *supra* note 40.

recruitment, hiring, employee assessment, benefits management, and other areas. While by no means an exhaustive list, the following are some compliance issues for employers and their counsel to consider:

- When using a third party to develop algorithms to identify potential candidates for a position, consider whether the broker could be considered a consumer reporting agency under the FCRA, thereby triggering the employer's obligations under FCRA.

- Consider whether any data inputs in algorithms used to evaluate individuals for hiring or recruiting purposes could be considered proxies for membership in a protected class or disproportionately have an adverse impact on a protected class.

- When using a third party to design an algorithm, evaluate the data inputs identified by the third party for potential disparate impact concerns.

- With respect to confidential employee health data collected in connection with a wellness program, obtain only de-identified and aggregate information from the wellness plan administrator and keep all wellness plan information in separate, secure files.

- Inquire about the security protocols that third-party vendors have in place to maintain the privacy and confidentiality of employee health data, and whether employee health data is sold to data brokers.

- Investigate whether a prospective vendor would be willing to offer an indemnification provision in its contract.

- When monitoring and analyzing employee activities inside and outside of the workplace, consider employees' reasonable expectations based upon existing company policies, and whether those policies need to be revised to put employees on notice of monitoring activities.

Guided by these and similar principles, employers and their counselors can navigate through uncharted territory while they wait for the law to respond to the challenges and opportunities presented by big data in the workplace.