Statement of Eric Dunleavy, Ph.D.
Director of the Personnel Selection and Litigation Support Services Group
DCI Consulting Group, Inc.
Washington, DC

On Behalf of the
Society for Human Resource Management

Presented to the
U.S. Equal Employment Opportunity Commission Hearing on
Use of Big Data in Employment
Wednesday, October 12, 2016
I would like to thank the Commission and the Society for Human Resource Management (SHRM) for providing me the opportunity to briefly discuss the current state of big data tools in HR practice, with particular focus on employment decision making. I present my perspective today as a SHRM member, an Industrial/Organizational (I/O) Psychologist, and as a scientist-practitioner who has conducted applied research and consulted for a wide range of employers across a variety of industries.

Importantly, I have no vested interest in whether employers use big data tools in employment decision making processes. My purpose in testifying today is to provide some contemporary perspective on what we currently know and don’t know about the use of big data tools in employment. As other panelists will likely share, this is a complex topic, and there is still little research on big data tools along a number of important dimensions for HR practitioners, and little precedent from EEO matters. My plan is to first describe big data and related tools in the context of employment, identify some HR-related research questions big data tools may be able to help answer, review how employers and HR practitioners specifically may use these tools, consider the role of human judgement, and conclude with some basic considerations related to equal employment opportunity.

**What do we mean by “big data” in employment?**

Big data can mean different things in different contexts. In the context of employment, we are discussing big data because technology has changed and streamlined the type and amount of data available to organizations. Big data can refer to a wide variety of tools typically characterized by (1) a large number of data points (referred to as volume), (2) information being added constantly and for immediate use (referred to as velocity) and (3) the integration of multiple sources (referred to as variety; Laney, 2001). Recent I/O work has documented the application of big data and novel analytic methods to the workplace and potential consequences on organizational science and practice (e.g., Tonidandel, King & Cortina, 2016).

Today, big data tools typically pair (1) large and evolving datasets from various sources of information with (2) machine learning methods (i.e., supervised and/or unsupervised learning approaches that identify rules from existing data to either predict an outcome or more appropriately operationalize it) that combine these sources of information in sophisticated ways (Illingworth, Lippestreu & Deprez Sims, 2016; Oswald & Putka, 2016).

**What questions are employers and HR trying to answer?**

Many questions employers are trying to answer via big data tools are not new. Employers have long been asking questions related to whether the future behavior of applicants or employees can be predicted from available information. Such individual outcomes often drive organizational outcomes. These research questions have existed in the I/O Psychology literature for many decades, particularly in the context of using biodata in employee selection. Biodata generally refers to a wide variety of work and life history data points, which have been used to predict important work outcomes like performance, absenteeism, employee attitudes and turnover (e.g., Cucina, Caputo, Thibodeaux, & MacLane, 2012; Mael, 1991). Whether it is deciding whom to hire, whom to promote, or how to develop training or other organizational interventions to influence attitudes and behavior, employers have long attempted to collect data to predict various work outcomes.
Similarly, HR has been successfully using data and HR metrics for decades. For this reason, SHRM has identified advocating for the use of metrics, analysis and evidence in making HR decisions as a core competency for HR professionals. The ability to interpret information with which to make business decisions and recommendations is key to informing many workforce questions. SHRM HR certification materials point out that “analytics have the potential to improve individual and organizational performance because they:

- Embed workforce intelligence as a cornerstone in management decision making.
- Improve workforce planning and forecasting.
- Shorten recruiting cycles.
- Reduce recruiting and separation costs.
- Retain critical talent.
- Drive succession planning.
- Use on-demand insights to avoid costly mistakes regarding the workforce.
- Redirect money spent on wrong employee initiatives to more beneficial programs.” (Society for Human Resource Management, 2015).

How are data being used?

As the above list shows, data can inform on a wide range of organizational dimensions. In my experience with clients and colleagues, I am most familiar with big data tools in the context of employee recruitment and hiring, where big data tools may be used to identify applicants and collect and evaluate information on applicants. Online recruiting focusing on external job boards, social media, and internal career platforms may be useful strategies for identifying talent. (Illingworth et. al, 2016). Applicants are easier to reach today, and can apply to many jobs from remote areas. In part because of this situation, automated steps at the front end of hiring processes may be particularly useful given the large size of internet-based applicant pools and the human capital effort required to evaluate those applicants on eligibility and qualifications.

Tools may take on a variety of forms. Some may resemble narrow application/resume reviews, others may look like broader biodata evaluations, still others may take the form of interviews, and others may not fit neatly into a traditional category. Sources of applicant data that could be used as part of a big data hiring tool may vary from relatively narrow self-report information submitted as part of an application or resume (e.g., work history, education, training, certifications) to broader personal data mined from social media sites or other publicly available information on the Internet. In other scenarios, video game play may be used to collect content on an applicant’s standing on various dimensions, or video interviews may be recorded and scraped for data related to what applicants say, how they say it, and what they look like saying it.

Data metrics for employees may also be available and leveraged to evaluate performance and/or make decisions related to promotion, placement or compensation. In these instances, objective data from Human Resource Information Systems (HRIS) and other organizational systems may be evaluated and combined. In each example described in this section, machine learning methods may be used to combine substantial amounts of information in ways intended to achieve organizational goals.
For additional details on broad HR applications of data analytics, I suggest that readers review “Use of Workforce Analytics for Competitive Advantage.” This report was published by The SHRM Foundation, in collaboration with the Economist Intelligence Unit (EIU), and formed the basis of a recently held Thought Leaders Retreat on the use of big data in employment.

What we know about contemporary use of big data

At the current time, little is known regarding how widely various HR big data practices are being used. SHRM recently conducted a survey of 279 members to better understand how HR professionals use big data (Kurtessis, Alonso, & Mulvey, 2015). I thought that the following was interesting:

- 32% of HR professionals reported that their organization uses big data to support HR; those in larger organizations (i.e., greater than 200 full time employees) were almost twice as likely to use big data tools as those in smaller organizations.

- Nearly all of the HR professionals surveyed (94%) indicated that their HR department was responsible for analyzing big data. Only 23% of respondents indicated that external vendors conducted analytics and 16% indicated that a non-HR function in the organization conducted analytics for HR.

- Of those who did not use big data to support HR, the most frequent reason (42%) was a lack of knowledge and expertise in compiling, analyzing, and interpreting data.

- 23% of the HR professionals surveyed indicated that data are analyzed with a specific business research question in mind; 9% of HR professionals indicated that they rely on data mining approaches to glean insights from their data without having a specific research question in mind; 68% indicated that they use a combination of data and research question driven approaches.

One other future-facing point is worth noting. In the recent report from The SHRM Foundation, a 2015 EIU survey was cited and reported that 82% of organizations plan to either begin or increase their use of big data in HR over the next three years. As such, it will be interesting to monitor the results of similar surveys in the years to come, where many changes may be likely in terms of HR demand, the nature of predictive analytics used, and the types of data that may be available.

The interplay of data and human judgment

One point of interest relates to the role of human judgment and expertise in big data tool development and use. One potential advantage of computer scoring and algorithm refinement is that at some point in the process human judgment and expertise in employment decision making may be reduced or even made unnecessary. As such, from an EEO perspective, traditional forms of intentional discrimination could eventually be minimized or removed from the decision making equation. However, at some point in the data gathering and tool development process, human judgment and expertise were likely involved to (1) determine an initial set of content inputs to evaluate, (2) select work outcomes to predict and prioritize, (3) identify particular prediction methods to combine information, and (4) determine whether adverse impact is of
interest to the employer, and if so, what groups are considered previously disadvantaged or not and how this issue is handled. These decision points are important to consider.

**Equal employment opportunity considerations**

As EEOC is aware, some have speculated that big data methods could have important implications for equal employment opportunity and diversity and inclusion outcomes, although little research is currently available (e.g., Ajunwa, Freidler, Scheidegger, & Venkatasubramanian, 2016; Barocas & Selbst, 2016; Morgan, Dunleavy, & Devries, 2016). Lawyers have already started to think about current regulatory frameworks and their application to big data tools, and SHRM and the broader employer community look forward to what legal scholars have to say on this issue. Below are a few practical notions that I thought may be worth sharing.

With regard to intentional discrimination via disparate treatment or pattern or practice, big data concerns may focus on whether direct protected group status membership (e.g., race, ethnicity, sex, national origin, age) is incorporated in some way in the algorithm used to evaluate content. If group membership is not included in the algorithm, and human judgment decisions on the front end did not directly consider protected group status, then intentional discrimination as a function of the algorithm may not be an issue of concern.

With regard to disparate impact, big data tools seem to fit with traditional Title VII scenarios and a “Uniform Guidelines style” approach to scrutiny. The first general question of interest may be whether the tool has substantial adverse impact on members of a protected group. If it does, the second question relates to job relatedness and business necessity, and traditional concepts of validation from the I/O Psychology realm may apply. A third consideration may be whether an equally valid less adverse alternative was available.

There has been some conversation around whether big data tools are truly selection procedures. The Society for Industrial and Organization Psychology (SIOP) defines such procedures as: an **assessment instrument used to inform a personnel decision such as hiring, promotion, or placement** (SIOP, 2003). This notion is generally consistent with the broader definition from the Uniform Guidelines on Employee Selection Procedures.\(^1\) If the tools are used to make high stakes employment decisions, they would seem to be no different than traditional selection procedures in this way.

A number of additional data complexities may need to be considered in the context of EEO, including:

- Ambiguity of what is being measured. The scope of data inputs and complexity of machine learning algorithms may make it difficult to identify exactly what variables are being considered and how they are being combined and used to make employment decisions.

- Constantly changing algorithms: The iterative nature of some algorithms may make isolating a specific procedure over time a challenge.

• Ethical and security issues, such as those involving confidentiality and privacy issues.

• Content inputs that may on their face violate laws (e.g., ADA, FMLA, FCRA, state pay laws).

Conclusions

Big data methods are being used in the employment setting. With those methods come potential risks and rewards. The prevalence of these methods is still unclear, and I reiterate that I have seen little empirical research related to the validity and adverse impact associated with big data tools used for employee selection. However, given the data analytic foundation of these tools, research is likely feasible in some cases. It may be prudent for big data tool users to consider the purpose of the tool, what is being measured, what research is available to support validation and use, and whether subgroup differences exist. I have been involved in situations where these factors were considered and big data tools were validated and implemented. I have also been involved in situations where employers chose not to implement big data tools after considering these issues.

The question of whether employers can leverage contemporary big data for employment decision making has been answered in the affirmative. Whether employers should do so, and how to go about it in their particular situation, are separate questions. It will be particularly interesting to monitor the research and practice literature over the coming years as more relevant information related to big data tools becomes available. We should all stay tuned.

References


