Title	Big data: lessons for employers and employees
Author(s)	Jeske, Debora; Calvard, Thomas
Publication date	2020-01-06
Original citation	Jeske, D. and Calvard, T. (2020) 'Big data: lessons for employers and employees', Employee Relations, 42(1), pp. 248-261. doi: 10.1108/ER-06-2018-0159
Type of publication	Article (peer-reviewed)
Link to publisher's version	http://dx.doi.org/10.1108/ER-06-2018-0159 Access to the full text of the published version may require a subscription.
Rights	**© 2019, Emerald Publishing Limited. This manuscript version is made available under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) licence.** https://creativecommons.org/licenses/by-nc/4.0/
Item downloaded from	http://hdl.handle.net/10468/9574

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Big data: Lessons for employees and employees

Debora Jeske
University College Cork

Thomas S. Calvard
University of Edinburgh

Corresponding author:

Thomas S. Calvard
Email: Thomas.Calvard@ed.ac.uk

Funding statement

The present research was not funded by any agency or grant.

Conflict of interest

The authors are not aware of any conflict of interest.

This manuscript was accepted by Prof. Nickson, editor for “Employee Relations”, on July 30th, 2019.

Big data: Lessons for employers and employees

Journal:	*Employee Relations*
Manuscript ID	ER-06-2018-0159.R2
Manuscript Type:	Research Paper
Keywords:	big data, consent, data analytics, employment, ethics

Big data: Lessons for employers and employees

Abstract

Purpose: The focus of the current article is to critically reflect on the pros and cons of using employee information in big data projects.

Approach: The authors reviewed papers in the area of big data that have immediate repercussions for the experiences of employees and employers.

Findings: The review of papers to date suggests that big data lessons based on employee data are still a relatively unknown area of employment literature. Particular attention is paid to discussion of employee rights, ethics, expectations, and the implications employer conduct has on employment relationships and prospective benefits of big data analytics at work for work.

Originality/value: This viewpoint article highlights the need for more discussion between employees and employers about the collection, use, storage and ownership of data in the workplace. A number of recommendations are put forward to support future data collection efforts in organisations.

Keywords

big data; consent; data analytics; employment; education; ethics

Introduction

Computer and internet-based technology has steadily advanced in recent decades and changed working lives, both positively (e.g., working flexibly, remotely and virtually) and negatively (e.g., work intensification, 24/7 availability). These technological developments are the source of a dramatic increase in the amount and availability of data in the world (McAfee and Brynjolfsson 2012). Larger sets of data can be captured more readily than ever before, increasing the potential for developing analytical formulae and rules to solve problems (via algorithms to process data), which in turn generate insights in the form of new information processing or decision-making aids (Dormehl 2014).

Big data has become a popular label for many data analytics efforts. Originally, the term big data emerged to define the technological revolution that enabled immense data collection (Jacobs 2009). Since then, the term has migrated into other domains and stands for different analytical aspects, depending on the context within which big data is mentioned. The term is now used to refer to both data processing capabilities and the characteristics of data, encapsulating both technical but also commercial aspects of data collection activities (Nunan and Di Domenico 2017). Mayer-Schönberger and Cukier (2013) consider big data as the emerging ability to crunch vast collections of information and analyse it instantly (see also Kitchin 2014). In a similar vein, boyd and Crawford (2012, pg. 663) suggest that big data is not necessarily a statement describing the size of data but instead a term that designates the "capacity to search, aggregate, and cross-reference large data sets."

The most important characteristic is the fact that big data analytics will go beyond traditional data sources (Ducey et al. 2015). Specifically, big data may involve several conjoined datasets from very different sources, and include data points generated from a variety of multimedia sources, such as video and audio records, pictures, different types of file formats captured presentations and texts, as well as sensors, frequencies, and behavioural traces such as clickstreams on websites (Zikopoulos et al. 2012). Within organisations, much of the employee-focused data may be gleaned from their use of social learning and collaboration tools on corporate intranets and social media platforms (Ducey et al. 2015). Many human resource information systems (HRIS) are data-rich repositories of information about employees, workplace events and conditions (Guzzo et al. 2015).

From a human resources (HR) perspective, big data represents a significant journey replete with opportunities and challenges. These include identifying and demonstrating novel sources of workforce value, evidence, and impact in ways not possible before, when HR was a back-office function, relatively untouched by strategic digital transformation and data revolution. Accordingly, the Chartered Institute of Personnel and Development (CIPD), a leading association for the HR profession in the UK, has produced a suite of reports and resources on big data, people analytics, talent analytics and related topics (e.g. CIPD 2013, 2017, 2018a, 2018b).

Big data is related to many ongoing issues and conditions to support its use. These include closing skills gaps, breaking out of silos to collaborate across the organisation, engaging with senior executive governance and decision-making, developing human capital metrics and data reporting standards, assessing people risks, positive relations of analytics cultures with business performance, and variable international adoption levels. Some commentators assert that small teams working on HR analytics are working through management cycles in various industries, like IT, aviation, and retail. After two or three years of having more data infrastructure and ecosystems in place, they can move past a ‘honeymoon period’ of initial learning toward the next stage of maturity in problem-solving with big data (Belizon 2019; Creelman 2019).

Pros and cons of big data for employers

At an organisational level, big data tools enable organisations to catalogue more personal communications than ever before, encompassing the views, sentiments, and behaviours of all those who interact with the organisational systems and interfaces, including employees (Lohr 2015). Big data has also allowed organisations to make improvements to existing processes (Nunan and Di Domenico 2017), often in the form of incremental progressions rather than significant reinventions. As with previous process improvement tools, the value of these small improvements cannot be underestimated, and the cost savings via improved organisational effectiveness can be significant. For example, organisations like Wal-Mart and Credit Suisse are using analytics to save turnover and churn costs by understanding who might quit the workforce and why, before they do so, saving millions a year where managers used the information anonymously to improve employee retention (Silverman and Waller 2015).

Nonetheless, there are persistent concerns revolving around poor data quality, and accountability, especially since employees are often active contributors to big data in organisations. Organisational accountability can remain unclear in relation to roles involved with keeping, processing, and analysing the data. In addition, many employees and managers may not agree with the extent to which employers seek confirmation in data rather than from employees themselves, increasing employee-focused data collection, but in the form of technological determinism rather than employee voice, and fostering the continuous quantification of HR practices. This argument is in line with concerns that data-driven modelling may also threaten the human agency associated with creativity, autonomy and self-determination (Dormehl 2014).

However, such concerns equally do not rule out potential opportunities for big data, AI models, and algorithms to take over mundane, repetitive tasks and free up and release more space and resources for employees to develop their soft skills and creativity in self-determining ways (Wilson et al. 2017). Using data systems to support real-time employee self-service and full integration and automation of leave, attendance, learning and performance development is another way of envisaging this.

Many of these data-driven tensions will involve working through the changes in terms of crucial constellations of organisational power (e.g. between union members and employers), workforce diversity, and access to employment (e.g. where big data is used to profile job candidates; Tonidandel et al. 2016). In the present manuscript, we focus mainly on new questions that arise for professionals tasked with big data analytics, evaluation and evidence-based implementation of findings of big data based on organisational (and specifically employee) records.

New data challenges for employers in the age of big data

Employers today are not only able to monitor employee activities on their computers and intranet but can also monitor their Internet activities via different devices, app activity, keystroke logging and other 'digital exhaust' (Harford 2014). They are also able to monitor employees during their offsite activities via their mobile devices, network traffic, and wearable devices (many of which employers will provide at no cost to their employees). In recent years, several researchers have started to identify ethics – in addition to privacy – as a major consideration in terms of how big data are generated, used, analysed and interpreted more generally (Gil de Zuniga and Diehl 2017). The following section will outline and discuss data-specific challenges as well as employee-related issues.

Quality, quantity and use

One of the key promises of big data is the knowledge that may be gained from it. While more data may not straightforwardly translate into larger amounts of better quality data, big data can mean improved access to good data or at least the existence of data where none previously existed (Stone et al. 2018). Whatever the merits of particular data, findings will be partly determined as they emerge from the positive and negative qualities of that data.

When more suboptimal, incomplete or poorer quality data are used, we can expect there to be problematic implications for the conclusions drawn. Overall, big data tends to be more unstructured than structured (Nunan and Di Domenico 2017), and a number of HR, Industrial-Organisational (I-O) psychologists and critical data scholars have expressed reservations about big data, emphasising recommendations for improving analytic practices (Angrave et al. 2016; boyd and Crawford 2012; Guzzo et al. 2015; Kitchin 2014). Metrics therefore need to be clearly defined and used uniformly in organisations, minimising and acknowledging essential details or qualifications left out (Lytle 2016; Roberts 2013). Analytics and algorithms will only be as good as the data put into them. Considerable time should be invested to evaluate and develop new algorithms with care (Arnaout 2012), in order to incorporate the work context within their metrics and analytic rules.

Data will vary in how fully they capture the information necessary to answer all employers' questions. First, some problem-solving may focus on issues that are the symptoms of, but not the origins or root causes of the presenting issue. Many workplace issues or problems underlying employment relations are multi-dimensional, comprised of elements that are difficult to model, even in big data, as these dimensions may be less tangible or unobservable in the workplace (e.g. psychological climates at work). This is where suboptimal or poor quality data can become a hurdle, in that it can provide some answers that an employer seeks, but will still be unlikely to tell a manager how to tackle specific employee, team and work process issues. As the metaphorical data 'haystack' containing the valuable needles - or insights – becomes more substantial, this merely increases the likelihood of identifying spurious trends (Aradau and Blanke 2015). Big data should therefore be carefully translated into meaningful action plans that could be utilised in the improvement of employment practices (Toterhi 2014). For example, Rasmussen and Ulrich's (2015) vignette detailed how Maersk drilling, an offshore drilling organisation, used data analytics to optimise employee performance, but only through a careful questioning and testing process relating links between team competence, error types, leadership quality and customer satisfaction.

Ultimately, the imperfect quality and quantity of big data can therefore be offset to some extent by more positive interpretive capabilities in how it is used by employers and employees (Calvard 2016). At the negative extreme, big data usage could involve mindless, careless, uninformed analyses that are more superficial and spurious in leading to adverse consequences. However, under more positive organisational conditions, big data could be used diligently and responsibly to make sense of complex phenomena from a range of perspectives.

Data storage and ownership

There are also methodological data storage and analytical challenges facing employers attempting to make sense of big data. Ducey et al. (2015, pg. 557) summarise this situation as follows: "Big data is not just about a large data set, it is asking us to filter petabytes of data per second from almost any connected device, analysing the data while still in motion, deciding what if any data must be stored, and even using analytics tools to virtually integrate the data with data stored in traditional warehouses." Additional challenges for employers involve the retrieval, storage, security and archival requirements for employee data – addressing these will often generate costs and require new procedures (Ducey et al. 2015).

Outsourcing data services to third-party vendors will change the dynamics of how employers share responsibilities for data analytics storage and ownership, as well as the sharing of resource-intensive data integration and conversion costs. External data analytics providers cannot entirely reduce employers' legal responsibilities towards safeguarding employee data.

As a professional function, HR therefore sits at the centre of a greater variety of data sources, but without having as much direct control over them. HR may be able to access some employee data from its own banks (e.g., payroll), but these sources may not be straightforwardly linked to other enterprise systems (Roberts 2013). Resolving these tensions through careful collaboration will support connecting internal datasets to external datasets, such as those made available by government agencies, as well as datasets containing industry benchmarks and workforce demographics (Roberts 2013). Much depends on the favourability of national conditions and the supply chain or stakeholder environment. Government and public sector big data use cases show many possibilities for strong partnerships; for example, weather patterns, law enforcement, health services, and regulatory compliance datasets are all publicly stored and could be used to solve workforce-related problems affecting productivity and well-being.

There will always be some risk of ambiguity or lack of shared responsibility for addressing the threats and fallout related to big data and data accidents (Nunan and Di Domenico, 2017). Right now, many stakeholders may not feel they have a stake in managing risks unless they have some personal stake in the data or have been directly affected by privacy violations. Becker and Smidt (2016) raise the point that exposure to risk requires that the parties involved actually have an interest in the outcome. Encouraging ownership of data but also the risks that arise from it will be vital to increasing accountability and forward thinking. Strategies for proactive accident handling and sector-wide knowledge sharing will also be needed for instances when accidents inevitably occur.

Ethics and employee data: A continuing debate

Employee monitoring is far from a new phenomenon in the world of work (Karim et al. 2015). However, the integration of data on a big data scale, and its use in predictive modelling, far exceeds previous analytical capabilities. We consider the ethical challenges facing organisations who use employee (big) data. In the age of big data, three central issues are at the forefront of concern: consent, different privacy expectations, and the lack of guidance for those tasked with data collection and analysis. We will discuss each of these briefly in turn.

Employees as data contributors

To date, data analytics, particularly big data analytics, raises new questions for how we respond to a situation where employees are themselves contributors to a multitude of interconnected datasets generated at work and away from work privately. Their role in generating this data raises critical questions regarding the treatment of employees as information sources and instruments in relation to data and algorithms. Other questions arise in terms of how employee interests in the employment relationship are being served or exploited, and how their actions are informing predictions about the future, predictions that may often be self-fulfilling. A related issue concerns data privacy and protection expectations that employees hold in regard to their employers, and how these may shape their acceptance or rejection of organisational data collection efforts. Some employees will react more positively than others to the prospect of carrying wearable devices provided by their employers, particularly invasive devices that register physiological responses, physical mobility, and attributes of vocalization (Guzzo et al. 2015). This brings us to several key points.

While consent may be a desirable and valuable component to any data collection effort, it is only likely to be a requirement in certain circumscribed situations involving designated study designs, participants, and research guidelines. When it comes to many other situations, such as the use of archival data or blind anonymous surveys, consent is less likely to be an explicit requirement. In the case of big data and workforce analytics, it would seem worthwhile to take a more careful look at the value of consent and informed consent where none may be explicitly required or expected – as is often the case over social media (Lam 2016). This is important given that employees are often unwilling participants in data collection efforts when the employer monitors their activities both on and off the job. The degree to which employees are aware of how their data was acquired and tracked by organisations (e.g., via cookies; Peacock 2014) often varies, with some more aware than others. In addition, any employee working online may not have wittingly given their informed consent to have their data used for organisational data analytics.

Consent thus remains a key concern with employee data, particularly if organisations assume that they have consent without necessarily ensuring for certain that consent has indeed been given in an informed, verified, and ongoing way (Custers 2016). In line with informed consent, users should be able to easily understand exactly which data they have given their consent to being collected and to obtain verification information when consent has been denied or annulled. However, in the case of big data, this will be extremely challenging to ensure in practice, given the complexity of changing patterns of consent into the data, and the potential use of anonymised data following the expiry of consent (Custers 2016).

Data triangulation poses similar dilemmas to consent. Triangulation is an inherent part of big data's 'big' appeal, reflecting valuable opportunities to verify and cross-validate insights across larger and more various sets of sources of data (Nunan and Di Domenico 2013). However, the risk of removing individuals' anonymity needs to be managed, where triangulated data connects information in ways allowing identification. This should require individuals' acknowledgement or permission to share, particularly if being used to make decisions affecting their situation (George et al. 2014).

In many cases, there are the data sets generated by employees due to their use of multiple tools and platforms. In essence, an employee contributes to multiple datasets simultaneously, a phenomenon that creates ideal conditions for triangulation. At present, however, employment laws do not consider the implications and threats that arise due to big data triangulation, so there are no safeguards in place to regulate triangulation, intentional or accidental. Similarly, methods of data analysis have emerged that allow employers to de-anonymize data sets. One such method is machine learning, a method that supports the automated analysis of analytical models and model building. New analytical tools such as these now allow employers to profile and identify individuals using social network information and HR details (Hashimoto et al. 2016).

A critical issue going forward is therefore to identify ways in which employers will ensure adequate levels of personal data control, supporting appropriate levels of self-determination and employee rights in the workplace. It is important to acknowledge that some big data projects can be experienced very positively by the workforce if they are democratically and transparently owned by the employees and used to further their own development. One such example is Google's project 'Oxygen' (Garvin 2013). Here, the software engineering workforce reported that using analytics to identify optimal management behaviours – as crucial as breathing oxygen, hence the name – was extremely positive and valuable because the data was 'about us, by us, and for us.'

Context still matters, even with big data

Big data analytics represent an opportunity for a more far-reaching analysis of employment contexts if the data is richer in quality and more thorough in its representation of a problem or phenomenon. However, these analytics also risk fostering greater complacency in neglecting the nuances of employment contexts if partial data is considered too readily to fully explain the context or trivialise its role (Johns 2018).

For instance, at present, most algorithms can only predict well-defined, short-term outcomes (Luca et al., 2016). Under conditions of uncertainty, and in the longer-term, such predictions may not be robust enough to warrant the trust placed in them. While longitudinal data collection (e.g. via live-streamed data) and analyses are increasingly possible, these are very resource-intensive to maintain and process, reducing incentives to contextualise analytics temporally. Importantly though, much of employee behaviour is a function of the employee's engagement with their contextual environment, at levels ranging from the general through to the more specific and local.

Furthermore, employees also engage in personal interactions with others in and outside the workplace, settings which may be far less predictable or controllable, and introduce unmeasured contextual variations and qualifications to prevailing interpretations. Advocates of big data analytics may assume that employees and their behaviour are observable, with fixed and highly generalisable characteristics. These assumptions of abstraction, reduction, and essentialism of context have implications for employees. Employees could be labelled by data (e.g. as opinion leaders or followers), for instance, given that past performance data is being used to semi-permanently label them as poor, average, or good performers, independent of other contextual factors, and outside of the original context of measurement. This categorisation may affect their inclusion in and access to future training and development opportunities.

The value of big data for employees lies in ensuring that it can provide or support more holistic, nuanced understandings of themselves and the practices undertaken in the context of their work. Big data decision-making therefore needs to be gradually embedded in governable routines in ways that do not sacrifice contextualisation by literally taking data-driven employee judgments too far 'out of context' (Janssen et al. 2017). One way to better take social context into account, for instance, is to use *relational analytics* that focus more on relationships between employees (influence, team networks, back-up support) instead of individual attributes in isolation. Most companies, with clear policies in place, can contextualise their analytics by collecting behavioural data from 'digital exhaust' – the emails, posts, logs, team formations, and project milestones recorded by digital applications and platforms (Leonardi and Contractor 2018).

Recommendations

In this section, we present several recommendations, not only for employers, but also for employees. We focus on improving ethically responsible conduct and education in order to avoid and recognise ethical workplace dilemmas, and the role of employers in improving systems design to respect privacy. We hope that our recommendations provide informative starting points for those tasked with managing the challenges for both employer and employees in the process of creating new data projects related to big data efforts.

Ethical data management via Chief Data/Information Officers

As implied above, ownership of or responsibility for data can remain elusive, for both the data analysts and the originators or subjects. This may be due to a lack of understanding by the data-generating users of what data is generated and who owns it (Foster 2014). Similarly, big data is often collected "regardless of, and potentially without knowledge of, the purpose for which it is to be finally used" (Nunan and Di Domenico 2017, p. 487). Levels of responsibility, accountability and awareness relate also to the recurrent expiration of licenses, products, and software updates that are run and abandoned in rapid-release cycles (Clark et al. 2014). In addition, the context surrounding how data are collected and analysed needs to be explored to reduce uncertainty over how these sets of data are being used (Nunan and Di Domenico 2017). These circumstances obfuscate the inherent privacy risks for the users who contribute to the data.

One option to counteract such situations is to create organisational (data) ethics boards (Medland 2016), or a similar kind of committee which would include chief data officers, employee and union representatives, where appropriate. In particular, functions like HR, legal compliance, and marketing will be increasingly likely to work closely together on establishing revamped data protection roles and responsibilities, such as chief data privacy officers with distinct expertise and independent decision-making powers (Smith, 2018). One example concerns how Google has attempted to set up an eight-person AI ethics board, called the Advanced Technology External Advisory Council, to ensure technology and data are ethically developed. It has had to scrap it within a week of its announcement, however, due to employee backlash over the political backgrounds and views of some of those appointed, highlighting the important of diverse employee representation in ethical data management initiatives (Price and Bastone 2019).

These issues now have increasing, wider relevance for many medium-sized and larger organisations that generate, use and analyse large datasets (or pay third parties to do so on their behalf). Good ethical practice and rigorous data security play an essential role in securing trust in organisations and strengthening their employee and stakeholder data management (e.g. Carucci 2016; Fung 2015). Analytics and algorithms can never achieve true 'objectivity', so it is important to challenge any strong truth-claims in organisational culture and management where appropriate (Aradau and Blanke 2015). In addition, such ethics boards or committees can identify when and where an employer should take steps to respect employee privacy and expectations in order to maintain productive and positive employment relationships. These boards may also include data scientists tasked with analysing big data sets (including data points collected from employees), and external experts to guide the process, in approaches similar to those used by research institutions to manage ethical research questions.

Ethics education for employers and employees

Analysts coming from the field of marketing, HR, or information systems are not necessarily trained in ethics, nor are they familiar with basic ethical procedures behind data collection. Many professionals who are asked to work with employee data are therefore facing a steep learning curve about ethics, as well as a lack of organisational incentives to encourage such engagement. Education, training, learning and development can together provide a means by which employers become more cognizant of the ethics of using big data. Exactly how employers decide to educate managers, employees and other stakeholders (e.g. union members) about their data efforts may be subject to particular organisational requirements and data use. However, what is key will be to establish ethical data management (as part of ethical leadership in the workplace or via the establishment of ethics boards).

While it is important to recognise the learning and development responsibilities of employers in setting ethical data analytics agendas from the top, the bottom-up responsibilities and ethical training needs of employees as users should be emphasised too. Individuals and collectives in the workplace have a more devolved responsibility to some extent for how they access and curate sensitive data in their roles, in managing their own health and well-being, for example (Moore and Piwek 2017). Employees should be widely encouraged to reflect on how they deploy privacy management strategies for themselves and others (Baruh and Popescu 2017), and to challenge unethical actions that seek to undermine or subvert the truth-value of data-driven systems and practices (De Laat 2018). The 'quantified self' movement emphasises how we collect and share big data about ourselves, and 'quantified employees' may be able to actively choose to anonymously share data they collect to hold management accountable through emerging websites and apps like Glassdoor (Bersin 2014).

Top-down and bottom-up forms of analytics education should include a discussion around and clear policy on the ownership of data, covering guidance on responsibilities for ownership, handling of data, and consent (Jeske and Shultz 2016). Stakeholder groups, including prospective employees, can learn about and jointly enforce the values and rules of data management ethics in conjunction with legal guidelines and data security imperatives (Jagger 2016). Publications by Angrave et al. (2016), boyd and Crawford (2012), Davenport et al. (2010), Guzzo et al. (2015), and Kitchin (2014) all represent highly recommended starting points for readers interested in learning more about practical applications and social and ethical issues surrounding big data in organisations and societies.

Managing data participation and privacy expectations

Dekas and McCune (2015) summarised several crucial questions for practitioners interested in using employee data in their data efforts. One of the questions asks if employers can expect employees to accept data collection as a form of research, and thus "assume [that] the research and application of findings are an expected or reasonable part of operating a business/organization" (Dekas and McCune 2015, p. 564). Even if confidentiality is assured and the inclusion of employee records presents no privacy risk to the employee, the basic question remains as to whether an employer can indeed justify the use of complex and detailed employee data as the best source of information to run a business effectively (Dekas and McCune 2015).

It may be judicious to start discussing employee expectations about what data is collected and shared. This also includes the degree to which employee data are employed, and for what purposes. Furthermore, clarity and expectations should be set in context around how likely it is that specific data will be used to inform decisions regarding the employment relationship (including retention, promotion, or pay rises). The use of data, their handling and role in the implementation of HR strategies and practices need to be clear from the start to avoid another 'black box' in HR.

Privacy and protection expectations will be influenced by how employers continue to make use of the data collected about individual employees (e.g. extending onsite data collection into offsite periods). Clearly Amazon's use of wireless wristbands to capture big data about its warehouse workers violated expectations about protecting workers' rights, but this prompted Amazon to respond by managing expectations that it would also be used to support employee well-being, not just productivity (Green 2018). Big data can itself also be part of the solution to managing employee expectations about other issues – for example, where the chief HR officer of IBM used social media data to detect employee dissent about a controversial travel ban and was able to communicate a reversal of the decision within 24 hours (Green 2018). Being clearer about the stages, durations, lifetimes, and different criteria of big data processing in employment are ways to break down the issue and unpack the nuances of different types of privacy expectation through more democratic and inclusive discussion (Chen and Yan 2016).

Purposeful design for data collection and use

Designing both data collection and data usage approaches with ethical purpose and adequate privacy are, we argue, important considerations for employers seeking to gain the most useful insights from big data; namely, to improve working conditions and address employee or organisational performance issues. The 2018 General Data Protection Regulation (GDPR) framework plays a pivotal role in many European workplaces, but we provide some key additional points here.

Any big data project in a workplace will involve decisions over the types of data considered, their collation, communication, and subsequent storage and usage, with the potential for ethical issues to arise at each step (Bazerman and Tenbrunsel, 2011). In addition, especially where algorithms, automation, and machine learning are involved, the data that organisations acquire from different devices and streams will require only minimal human agency or intervention.

This has additional ethical implications for the personal freedom and integrity of those stakeholders yielding the data (Fung 2015; Zwitter 2014). It may not even be necessary to force continuous monitoring of employees. Big data analytics can enable predictions to be made about employees who do not give consent (Custers, 2016), by using equivalent information from those who gave consent and who share at least some similar characteristics, according to a variety of data sources (Kosinski et al. 2013).

Nevertheless, some restrictions on data collection are within the control of the employer and can be used to ensure more positive reactions to big data initiatives. Privacy protocols can act to anonymise employees by stripping identifying information from devices. This needs to be balanced against the fact that if the data is stripped of too many identifying marks the functionality of the devices is affected, and utility for big data analytics will be greatly reduced. The combination of datasets may still enable analysts to de-anonymise them, despite these precautions (Nunan and Di Domenico 2017), as in the case of triangulation. Overall, this means that designing resilient systems with transparent and reasonable purposes and for privacy should be a vital to managing employment relations, employee expectations, commitment to the employer and the psychological contract. By implication, this suggests greater recognition of trade-offs among design criteria, such that the extensiveness of data collection via different devices and means may exist in tension with the attempts to design for privacy.

Beyond design for privacy, there is also the intention behind the combination of datasets that needs to be fit for purpose. In sum, device functionality, the purpose of data collection, and personal rights need to be carefully balanced (Li et al. 2016). Accordingly, authors like Guzzo et al. (2015) have produced further research guidelines for working with big data. These focus on consent, privacy and personal information (protection strategies to maintain data privacy), but also the need to consider issues regarding the actual integrity of data and data analysis (e.g. reporting on sampling issues, measurement and data quality, and minimization of potential harm).

Big data collection cannot and perhaps should not be unduly restricted, but where there is ample data access and collection occurring in employment settings, responsible usage should be encouraged and irresponsible misuse discouraged, beyond its collection. The work of Tom Davenport (e.g. Davenport et al. 2010; Davenport and Kudyba 2016) is particularly helpful on the various uses, abuses, and mistakes involved in using analytics and data-driven approaches to frame important insights and decisions. Responsible usage tends to revolve around normative considerations, such as fairness, inclusion, transparency, and the need for robust and flexible interpretation. Using data analytics for carefully designed and responsibly limited purposes flows from the data collection phase, but also takes people analytics away from being more of a fad and more toward being a genuinely innovative and valuable technique (Rasmussen and Ulrich 2015).

One possibility is that big data analytics can generate a provocative shock or serve as a catalyst for more innovative modes of knowledge production in and about workplaces, counteracting the flaws of more traditional uses of data, such as deductive hypothesis-testing and excessive theory-building. Potentially, less *deductive* attention would be paid to cumbersome issues of theories and methodologies, and more *inductive* attention to open exploration of puzzling phenomena and drawing equally open, unbiased conclusions (Chiaburu 2016; McAbee et al. 2017).

Embracing ethical big data leadership – A final comment

We wish to conclude with one final recommendation that we feel should become a core component of leadership in business, particularly for those using big data analytics. This is the assertion that ethical leadership in terms of data collection and management will be essential to informing and guiding employees and managers in the here and now, as well as in the future. To date, Silicon Valley ethics and leadership failures are well-documented, and are arguably not setting the bar high enough for other organisations and employers increasingly engaging in big data (Gobble 2018).

Our data-related decisions, considering current and expanding storage capacities, could have major implications for the lifelong working experience of employees, where there are new requirements to support data transportability rights - according to the 2018 General Data Protection Regulations, for example. Having an ethical role model has been shown to positively contribute to higher levels of subordinate-rated ethical leadership in organisational behaviour (Brown and Trevino 2014). Ethical leadership itself has been defined by Brown and Trevino (2006), both in terms of leaders being morally guided in their general interaction with others (a moral person), as well as behaving ethically in the context of working alongside others (a moral manager).

Ethical leadership thus requires not only fair and principled decision-making, but also the role modelling of ethical standards, the communication of ethical standards and holding those infringing on ethical standards accountable for their actions. Many of these behaviours will play an important role when employers use data analytics that include large datasets based on employees' interactions, social network engagements, and behavioural traces. Ethical leadership has also been positively linked to the consideration of employees, trust in leaders, honesty, and interactional justice (Brown, Trevino, and Harrison 2005). Embracing ethical leadership skills, therefore, has become highly significant and relevant as data collection at work becomes more prevalent, and the need for data to work for the benefit of employees as well as employers becomes more urgent.

References

Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M. and Stuart, M. (2016), "HR and analytics: Why HR is set to fail the big data challenge", *Human Resource Management Journal,* Vol. 26 No. 1, pp. 1-11.

Aradau, C. and Blanke, T. (2015), "The (Big) Data-security assemblage: Knowledge and critique", *Big Data & Society,* Vol. 2015, No. July–December, pp. 1-12.

Arnaout, R. (2012), "Elementary, my dear Doctor Watson", *Clinical Chemistry,* Vol. 58 No. 6, pp. 986–988.

Baruh, L. and Popescu, M. (2017), "Big data analytics and the limits of privacy self-management", *New Media & Society,* Vol. 19 No. 4, pp. 579-596.

Bazerman, M. H. and Tenbrunsel, A.E. (2011), "Ethical breakdowns", *Harvard Business Review*, April 2011. Available at: https://hbr.org/2011/04/ethical-breakdowns (accessed April 14, 2019).

Becker, K. and Smidt, M. (2016), "A risk perspective on human resource management: A review and directions for future research", *Human Resource Management Review,* Vol. 26 No. 2, pp. 149-165.

Belizon, M. (2019), "The HR analytics management cycle", *Analytics in HR (AIHR).* Available at https://www.analyticsinhr.com/blog/the-hr-analytics-management-cycle/?utm_source=activecampaign&utm_medium=email&utm_campaign=newsletter&utm_content=weekly-newsletter (accessed April 1, 2019)

Bersin, J. (2014), "Quantified self: Meet the quantified employee. *Forbes.* Available at: https://www.forbes.com/sites/joshbersin/2014/06/25/quantified-self-meet-the-quantified-employee/#be370fcc5fe4 (accessed 4 April, 2019) boyd, d. and Crawford, K. (2012) , "Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon", *Information, Communication & Society,* Vol. 15 No. 5, pp. 662-679.

Brown, M. E. and Trevino, L. K. (2006), "Ethical leadership: A review and future directions", *Leadership Quarterly,* Vol. 17 No. 6, pp. 595–616.

Brown, M. E., Trevino, L. K. and Harrison, D. A. (2005), "Ethical leadership: A social learning perspective for construct development and testing", *Organizational Behavior and Human Decision Processes,* Vol. 97 No. 2, pp. 117–134.

Brown, M. E. and Trevino, L. K. (2014), "Do role models matter? An investigation of role modelling as an antecedent of perceived ethical leadership", *Journal of Business Ethics,* Vol. 122 No. 4, pp. 587-598.

Calvard, T. S. (2016), "Big data, organizational learning, and sensemaking: Theorizing interpretive challenges under conditions of dynamic complexity", *Management Learning,* Vol. 47 No. 1, pp. 65-82.

Carucci, R. (2016), "Will your ethics hold up under pressure?", *Forbes*, 3 February. Available at: http://onforb.es/1SsLCOf (accessed April 13, 2016).

Chen, H. and Yan, Z. (2016, November), "Security and privacy in big data lifetime: A review", in *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage* (pp. 3-15). Springer, Cham.

Chiaburu, D. S. (2016), "Analytics: A catalyst for stagnant science?", *Journal of Management Inquiry,* Vol. 25 No. 1, pp. 111-115.

CIPD (2013), *Talent analytics and big data – the challenge for HR.* Research report. November 13. Available at https://www.cipd.co.uk/Images/talent-analytics-and-big-data_2013-challenge-for-hr_tcm18-9289.pdf (accessed April 1, 2019)

CIPD (2017), *Human capital metrics and analytics.* Technical report. Available at https://www.cipd.co.uk/Images/human-capital-metrics-and-analytics-assessing-the-evidence_tcm18-22291.pdf (accessed April 1, 2019)

CIPD (2018a), *People analytics: Driving business performance with people data.* Global research report. Available at https://www.cipd.co.uk/Images/people-analytics-report_tcm18-43755.pdf (accessed April 1, 2019)

CIPD (2018b), *Hidden figures: How workforce data is missing from corporate reports.* Report. Available at: https://www.cipd.co.uk/Images/hidden-figures_2018-how-workforce-data-is-missing-from-corporate-reports-executive-summary_tcm18-45152.pdf (accessed April 1, 2019)

Clark, S., Collis, M., Blaze, M. and Smith, J. M. (2014), "Moving Targets: Security and Rapid-Release in Firefox", *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security* (CCS'14), November 3-7, Scottsdale, Arizona, USA.

Creelman, D. (2019), "People analytics, is the honeymoon over? Online article. *Analytics in HR (AIHR).* Available at: https://www.analyticsinhr.com/blog/people-analytics-is-the-honeymoon-over/?utm_source=activecampaign&utm_medium=email&utm_campaign=newsletter&utm_content=weekly-newsletter (accessed April 1, 2019)

Custers, B. (2016), "Click here to consent forever: Expiry dates for informed consent", *Big Data & Society*, Vol. 2016 No. January–June, pp. 1-6.

Davenport, T. H., Harris, J. and Shapiro, J. (2010), "Competing on talent analytics", *Harvard Business Review,* Vol. 88 No. 10, pp. 52–58.

Dekas, K. and McCune, E. A. (2015), "Conducting ethical research with big and small data: Key questions for practitioners", *Industrial and Organizational Psychology,* Vol. 8 No. 4, pp. 563-567.

Davenport, T. H. and Kudyba, S. (2016), "Designing and developing analytics-based data products", *MIT Sloan Management Review,* Vol. 58 No. 1, pp. 83.

De Laat, P. B. (2018), "Algorithmic decision-making based on machine learning from Big Data: Can transparency restore accountability? *Philosophy & Technology,* Vol. 31 No. 4, pp. 525-541.

Dormehl, L. (2014), *The Formula: How Algorithms Solve All Our Problems – And Create More.* New York: Penguin.

Ducey, A. J., Guenole, N., Weiner, S. P., Herleman, H. A., Gibby, R. E. and Delany, T. (2015), "I-Os in the vanguard of big data analytics and privacy", *Industrial and Organizational Psychology,* Vol. 8 No. 4, pp. 555-563.

Foster, A. (2014), "The game is changing: The Business Information Survey 2014", *Business Information Review*, Vol. 31 No. 1, pp. 14-40.

Fung, K. (2015), "The ethics conversation we're not having about data", *Harvard Business Review*, 12 November 2015. https://hbr.org/2015/11/the-ethics-conversation-were-not-having-about-data (accessed April 13, 2016).

Garvin, D. (2013), "How Google sold its engineers on management", *Harvard Business Review, December.* Online article. Available at https://hbr.org/2013/12/how-google-sold-its-engineers-on-management (accessed April 2, 2019)

George, G., Haas, M. R. and Pentland, A. (2014), "Big data and management", *Academy of Management Journal,* Vol. 57 No. 2, pp. 321-326.

Gil de Zuniga, H. and Diehl, T. (2017), „Citizenship, social media, and big data: Current and Future Research in the Social Sciences", *Social Science Computer Review*, Vol. 35 No. 1, pp. 3-9.

Gobble, M. M. (2018), "Rethinking the Silicon Valley myth", *Research-Technology Management,* Vol. 61 No. 1, pp. 64-67.

Green, D. (2018), “Don’t forget the ‘H’ in HR: Ethics and people analytics”, *LinkedIn.* Online article. Available at: https://www.linkedin.com/pulse/dont-forget-h-hr-ethics-people-analytics-david-green/ (accessed April 5, 2019)

Guzzo, R. A., Fink, A. A., King, E., Tonidandel, S. and Landis, R. S. (2015), “Big data recommendations for Industrial-Organizational Psychology”, *Industrial and Organizational Psychology,* Vol. 8 No. 4, pp. 491-508.

Harford, T. (2014),” Big data: A big mistake?”, *Significance* Vol. 11 No. 5, pp. 14-19.

Hashimoto, E., Ichino, M., Kuboyama, T., Echizen, I. and Yoshiura, H. (2016), “Breaking anonymity of social network accounts by using coordinated and extensible classifiers based on machine learning”, In Y. Dwivedi et al. (eds). *Social Media: The Good, the Bad, and the Ugly*. I3E 2016 (pp. 455-470). Lecture Notes in Computer Science, Vol. 9844. Springer.

Jacobs, A. (2009), “Pathologies of Big Data”, *Communications of the ACM,* Vol. 52 No. 8, pp. 36-44.

Jagger, S. (2016), “Rules and Values: Expectations for a Business Ethics course”, *Chartered Association of Business Schools, Newsletter* 13 April 2016.

Janssen, M., van der Voort, H. and Wahyudi, A. (2017), “Factors influencing big data decision-making quality”, *Journal of Business Research,* Vol. 70 No. January, pp. 338-345.

Jeske, D. and Shultz, K. S. (2016), “Using social media content for screening in recruitment and selection: pros and cons”, *Work, Employment and Society,* Vol. 30 No. 3, pp. 535-546.

Johns, G. (2018), “Advances in the treatment of context in organizational research”, *Annual Review of Organizational Psychology and Organizational Behavior,* Vol. 5 No. January, pp. 21-46.

Karim, M. N., Willford, J. C. and Behrend, T. S. (2015), "Big data, little individual: Considering the human side of big data", *Industrial and Organizational Psychology,* Vol. 8 No. 4, pp. 527-533.

Kitchin, R. (2014), "Big Data, new epistemologies and paradigm shifts", *Big Data & Society,* Vol. 2014 No. April-June, pp. 1-12.

Kosinski, M., Stillwell, D. and Graepel, T. (2013), "Private traits and attributes are predictable from digital records of human behavior*", Proceedings of the National Academy of Sciences of the United States of America,* Vol. 110 No. 15, pp. 5802–5805.

Lam, H. (2016), "Social media dilemmas in the employment context", *Employee Relations,* Vol. 38 No. 3, pp. 420-437.

Leonardi, P. and Contractor, N. (2018). "Better people analytics", *Harvard Business Review, November-December.* Online article. Available at: https://hbr.org/2018/11/better-people-analytics (accessed 2 April, 2019)

Li, S., Tryfonas, T. and Li, H. (2016), "The Internet of Things: a security point of view", *Internet Research,* Vol. 26 No. 2, pp. 337-359.

Lohr, S. (2015), *Data-ism: The Revolution Transforming Decision Making, Consumer Behavior, and Almost Everything Else.* London: OneWorld Publications.

Luca, M., Kleinberg, J. and Mullainathan, S. (2016), "Algorithms need managers, too. Know how to get the most out of your predictive tools", *Harvard Business Review*, Vol. 2016 No. January-February, pp. 96-101.

Lytle, T. (Feb 2016), "Help wanted: HR analysts", *HRMagazine*, Vol. 61 No. 1, pp. 32, 36, 38.

Mayer-Schönberger, V. and Cukier, K. (2012), *Big Data: A Revolution that will Transform how we Live, Work, and Think*. New York, New York: Houghton Mifflin Harcourt.

McAbee, S. T., Landis, R. S. and Burke, M. I. (2017), "Inductive reasoning: The promise of big data", *Human Resource Management Review,* Vol. 27 No. 2, pp. 277-290.

McAfee, A. and Brynjolfsson, E. (2012), “Big data: The management revolution”, *Harvard Business Review,* Vol. 90 No. 10, pp. 60–68.

Medland, D. (2016), “Taking Ethics Seriously: By Setting Up Board Committees?”, *Forbes*, Vol. 9 No. March. Available at: http://onforb.es/1Tse7f6 (accessed April 13, 2016).

Moore, P. and Piwek, L. (2017), “Regulating wellbeing in the brave new quantified workplace”, *Employee Relations,* Vol. 39 No. 3, pp. 308-316.

Nunan, D. and Di Domenico, M. (2013), “Market research and the ethics of big data”, *International Journal of Market Research,* Vol. 55 No. 4, pp. 505-520.

Nunan, D. and Di Domenico, M. (2017), “Big data: A normal accident waiting to happen?”, *Journal of Business Ethics*, Vol. 145, No. 3, pp. 481-491. doi: 10.1007/s10551-015-2904-x

Peacock, S. E. (2014), “How web tracking changes user agency in the age of Big Data: The used user”, *Big Data & Society*, Vol. 2014, No. July–December, pp. 1–11.

Price, R. and Bastone, N. (2019), “After an employee backlash, Google has cancelled its AI ethics board a little more than a week after announcing it”, *Business Insider.* Available at: https://amp.businessinsider.com/google-cancels-ai-ethics-board-2019-4 (accessed April 5, 2019)

Rasmussen, T. and Ulrich, D. (2015), “Learning from practice: how HR analytics avoids being a management fad”, *Organizational Dynamics,* Vol. 44 No. 3, pp. 236-242.

Roberts, B. (2013), “The benefits of big data”, *SHRM HR Magazine*, 10 January. Available at: https://www.shrm.org/publications/hrmagazine/editorialcontent/2013/1013/pages/1013-big-data.aspx (accessed May 23, 2016).

Silverman, R. and Waller, N. (2015), “The algorithm that tells the boss who might quit”, *Wall Street Journal*. Available at https://www.wsj.com/articles/the-algorithm-that-tells-the-boss-who-might-quit-1426287935 (accessed March 25, 2019)

Smith, A. (2018), "GDPR says companies must have a data privacy officer: Who should that be?", *SHRM* online article, 23 May. Available at: https://www.shrm.org/resourcesandtools/legal-and-compliance/employment-law/pages/global-gdpr-data-protection-officer.aspx (accessed April 5, 2019).

Stone, C. B., Neely, A. R. and Lengnick-Hall, M. L. (2018), "Human Resource Management in the Digital Age: Big Data, HR Analytics and Artificial Intelligence". In P. Novo Melo and C. Machado (Eds), *Management and Technological Challenges in the Digital Age* (pp. 13-42). Boca Raton, FL: CRC Press.

Toterhi, T. (2014), "Make sure big data adds up: the science of data collection is meaningless without the art of interpretation", *Training,* Vol. 51 No. 4, pp. 14.

Wilson, H. J., Daugherty, P. and Bianzino, N. (2017), "The jobs that artificial intelligence will create", *MIT Sloan Management Review,* Vol. 58 No. 4, pp. 14-16.

Zikopoulos, P. C., Eaton, C., deRoos, D., Deutsch, T. and Lapis, G. (2012), *Understanding big data: Analytics for enterprise class Hadoop and streaming data*. New York, NY: McGraw-Hill.

Zwitter, A. (2014), "Big Data ethics", *Big Data & Society,* Vol. 2014, No. July-December, pp. 1-6.