

- Westin, A. F. (1967). *Privacy and freedom*. New York, NY: Atheneum.
- Zikopoulos, P. C., Eaton, C., deRoos, D., Deutsch, T., & Lapis, G. (2012). *Understanding big data: Analytics for enterprise class Hadoop and streaming data*. New York, NY: McGraw-Hill.

## Conducting Ethical Research With Big and Small Data: Key Questions for Practitioners

Kathryn Dekas

*Google, Mountain View, California*

Elizabeth A. McCune

*Microsoft, Redmond, Washington*

The focal article (Guzzo, Fink, King, Tonidandel, & Landis, 2015) sought to “raise awareness and provide direction with regard to issues and complications uniquely associated with the advent of big data” (p. 492), and we commend their success in offering Society for Industrial and Organizational Psychology (SIOP) members a solid foundation and resources on which to draw. Our aim here is to extend their position, particularly to drive the conversation toward concrete recommendations for how industrial and organizational psychologists (I-Os) working in industry can apply the principles set forth in the focal article in our day-to-day work, specifically around the issue of avoiding ethical missteps in this new landscape.

Our ideas described below are the product of a working group assembled prior to the SIOP 2015 conference in preparation for a panel discussion titled “Guidelines for Ethical Research in the Age of Big Data” (McCune et al., 2015). The working group included four I-Os working in the tech, retail, and consumer product goods industries; an employee data privacy expert from the tech industry; an associate director of an institutional review board (IRB) at a top university; and a member of a European Works Council.

The original aim for the panel was to provide SIOP session attendees with the proverbial “dos” and “don’ts” list in conducting ethical research with big data to help newcomers to the big data/data science world engage with these new methods in an ethically sound way. However, as we worked through the process it became increasingly clear that issues of ethics around the use of data—big or small—are highly subjective and context dependent,

Kathryn Dekas, Google, Mountain View, California; Elizabeth A. McCune, Microsoft, Redmond, Washington.

Correspondence concerning this article should be addressed to Kathryn Dekas, Google, 1600 Amphitheatre Parkway, Mountain View, CA 94043. E-mail: [kdekas@google.com](mailto:kdekas@google.com)

as is often the case in questions of ethics beyond a certain set of fairly non-negotiable standards. As pointed out in the focal article, even achieving these nonnegotiable standards is often not enough to assure that your research will be viewed as ethical by participants and others. As a result, our working group ended up asking *questions* more often than we were able to provide *answers*. These questions prompted meaningful dialogue and helped to surface unstated assumptions. We came to the conclusion that posing a set of core questions for researchers to consider within their own research context would be most useful in setting folks on the right path.

Our working group converged on the following general questions for researchers to ask themselves as they engage in a project involving big data—or, ideally, any research project. Although we focus on big data here because that is the topic at hand, the questions we pose really apply to all data/research, big or small. We advise asking these questions early and often, to help inform and govern decisions related to the study, and following up with more specific questions as the project unfolds. It is unlikely that any of these questions will have a clear “yes” or “no” answer, but seeking to answer them to the best of your ability, ideally with input from employees themselves and your employee privacy counsel, can shed light on whether your research is likely to generate ethical concerns.

First, **how comfortable is your employee population with the use of data and analytics in people-related questions/decisions?** General comfort with data and their use varies greatly across populations. In the tech industry, for example, employees tend to be familiar with the use of data mining and predictive analytics in product development and general business practice, which may translate into a greater comfort with using similar data and analytical techniques in internal research to inform people-related decisions. But, even within an industry that tends to have comfort and familiarity with complex data and analysis, employees may take a different view when the topics, data, and analysis are specific to the workplace.

Similarly, **would employees assume the research and application of findings are an expected or reasonable part of operating a business/organization?** The focal article raised the point that American Psychological Association’s (2010) code of ethics indicates researchers may dispense with informed consent if the research involves “the study of factors related to job or organization effectiveness conducted in organizational settings for which there is no risk to participants’ employability, and confidentiality is protected.” Although this is the American Psychological Association party line, ethical issues may still arise from employees even if an official body has deemed (or would deem) the research squarely related to organization effectiveness and with no risk to employees. It is prudent to take a conservative look at the research from the standpoint of research participants and

consider whether the average employee in your organization would agree that the organization *should be* conducting research of this type in order to best run its business most effectively.

Third, upon hearing about the research, **would employees feel like the researchers and the organization have their best interests in mind?**

The answer to this question is likely to relate to the psychological contract (Rousseau, 1995) and perceived level of trust between employees within the organization, as well as between employees and management. If the requisite level of trust between employee and employer is not present, employees may have a range of concerns about organizational research efforts (big data or otherwise) that could include inappropriate use of their personal data or survey responses, as well as concerns around how the results of the research will be used.

Similarly, **would employees feel a sense of violation if they learned the details about the study or its conclusions?** Asking this question may help put yourself in the employee's shoes. It hopefully goes without saying that if the answer is "yes" or "maybe" to this question, it is best to reconsider the goals or process for the study.

Finally, **can employees trust the analysis process to be objective?** This is perhaps the question over which the researcher has most control. Employees should have confidence that results will not be analyzed or interpreted in a biased fashion in order to achieve a more desirable result. It is incumbent on researchers to hold themselves to this high standard and also to be as transparent as possible with employees about results, no matter whether results are favorable.

### **Building the Partnership Between Researcher–Practitioners and Employees**

Conducting research within an organization of which you are a part presents unique opportunities (i.e., access to data and interesting questions) but also responsibilities. When your aim is to build and maintain a healthy partnership between researchers and employees who are aligned around organizational aims, these ethical considerations are paramount. It is always important to land on the side of ethical research, and conducting research within the organization of which you are a part introduces extra motivation to do so. It is important to get the above questions right. What can you do if you answer unfavorably to these questions now? We recommend three things:

First, **build a culture of data-driven employee practices.** It is unlikely employees will feel comfortable with advanced data and research techniques, such as those related to the use of big data, if more basic data/analytics are not already commonly used or socialized. Start small: Conduct basic analyses and communicate the results to employees, and emphasize in broad communications and training activities (e.g., new hire orientation) that the

organization values the inclusion of science and data in its people practices. Use any opportunity you have (e.g., launch of new employee programs, communication of compensation/benefits changes, hiring activities) to infuse language around the inclusion of data and analysis to inform decision making so employees come to understand the philosophy with which the organization manages people.

Next, **embrace transparency and value data privacy**. In communicating research results, but also more broadly, transparency helps build trust between employees and between leaders and employees. This trust is critical in giving employees confidence to assume good intent about business practices and research activities. However, it is also important to give employees confidence in your commitment to data privacy and handling their data with care. It is likely you will be able to engage in more complex research activities if employees trust from experience that their data are going to be used in a way with which they are comfortable.

Another suggestion is to discuss the potential ethical concerns of a research project with employees before beginning the project. Very few projects will not present at least some ethical questions. Rather than guess, potentially incorrectly, at how these issues will be perceived or received by employees, it is easier to go to the source. Putting together a group of trusted employees who have a good pulse on the tone of the broader employee base can help avoid unintentional missteps along the way. For some organizations, creating a kind of internal IRB or research ethics committee comprising an I-O or other social researcher, a member of the human resources (HR) team, and a member of the legal team, for example, may be another good option.

Finally, find ways to **educate your organization about research ethics**. Education across the organization not only may contribute to building a culture of data-driven employee practices but also will help establish common ground between researchers and employees with respect to expectations around ethical research. Consider who in your organization needs to be educated (e.g., research team members and other potential users of employee data, HR clients, typical employees/prospective research participants), the learning objectives you would have for each audience, and your strategy for getting these groups from their current level of knowledge to where they need to be.

In sum, we are grateful to the authors of the focal article for providing a very solid foundation for the SIOP membership about big data, its use by I-O psychologists, and the multitude of ethical issues that it brings to the forefront. We hope that our commentary will help practitioners better navigate the complex ethical issues inherent in this kind of research and that in the future we see more projects utilizing these types of data and methods.

## References

- American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct*. Washington, DC: American Psychological Association.
- Guzzo, R. A., Fink, A. A., King, E., Tonidandel, S., & Landis, R. S. (2015). Big data recommendations for industrial–organizational psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 8(4), 491–508.
- McCune, E. A., Dekas, K., Anderson, A., Kasmatis Singleton, M., MacNiven, S., & Sinnett, S. (2015, April). *Guidelines for ethical research in the age of big data*. Panel presented at the 30th Annual Conference of the Society for Industrial and Organizational Psychology, Philadelphia, PA.
- Rousseau, D. M. (1995). *Psychological contracts in organizations: Understanding written and unwritten agreements*. Thousand Oaks, CA: Sage.

# Big Data in I-O Psychology: Privacy Considerations and Discriminatory Algorithms

A. James Illingworth

*Geode People, Inc., Decatur, Georgia*

## Introduction

The “big data” movement is forcing many fields to establish best practices for the collection, analysis, and application of big data, and the field of industrial–organizational (I-O) psychology is not exempt from this disruptive influence. Over the last several years, I-O scientists and practitioners have grappled with questions related to the definition, application, and interpretation of big data (e.g., Doverspike, 2013; Maurath, 2014; Morrison & Abraham, 2015; Poepelman, Blacksmith, & Yang, 2013). The focal article by Guzzo, Fink, King, Tonidandel, and Landis (2015) continues this discussion and represents one of the first attempts to establish a formal set of recommendations for working with big data in ways that are consistent with I-O psychology’s professional guidelines and ethics requirements.

The big data issues discussed by Guzzo et al. are not unique to I-O psychology. In fact, they overlap significantly with similar discussions occurring among computer scientists, technologists, privacy advocates, and policymakers about the challenges of maintaining privacy, informed consent, and analytical rigor in the big data era. That so many other fields are engaged in a similar conversation provides a tremendous opportunity for I-O psychology to draw on insights from this larger dialogue to shorten the big data learning curve, ensure alignment with current thought in other fields,

A. James Illingworth, Geode People, Inc., Decatur, Georgia.

Correspondence concerning this article should be addressed to A. James Illingworth, 2955 Fantasy Lane, Decatur, GA 30033. E-mail: [ajillingworth@geodepeople.com](mailto:ajillingworth@geodepeople.com)