The Transformational Impacts of Big Data and Analytics

By Christy Walters & Crystal Walters

Introduction
The age of big data and analytics is changing the business and academic landscape as today’s organizations are using innovative techniques to derive valuable insights from unstructured and structured data. There is an evolving positive view of big data as its transformational impacts on business has come to the forefront in recent years. In this article, we will evaluate the current literature regarding big data and analytics and its transformative power. We will discuss the innovative ways in data analytics has evolved and how it is transforming today’s business landscape. Finally, we will form a conclusion on the impacts of big data and analytics on present-day information management practice and address the skills that are required for information management practitioners to create value for organizations in this big data age. The positive outcomes of big data and analytics do not come without emerging challenges that will also be considered in the review.

The ‘Big Data’ Evolution
In his 2012 article, “On the Origins and Development of ‘Big Data’,” Francis Diebold concludes that the phrase ‘big data’ likely arose out of lunch-table conversations involving Chief Scientist, John Mashey at Silicon Graphics Inc. (SGI) in the mid-1990s. In a later interview, Mashey stated his first uses of the term stemmed from his desire to convey in the shortest way that “the boundaries of computing keep advancing.”

The concept of big data was first defined in terms of the 3Vs by Doug Laney in his research note “3D Data Management: Controlling Data Volume, Velocity, and Variety” in
2001. Two decades later, the 3Vs of volume, velocity and variety are the generally accepted defining dimensions of big data and have evolved into the 4Vs with the addition of the V of veracity. The 4Vs have provided a framework for the design of the innovatory software required to handle the needs of the big data explosion. In the current information age, big data and computer analytical techniques are used to process and glean insights from applications that are so large (from terabytes to zettabytes) and complex that they require advanced and unique data storage, management, analysis, and visualization technologies.

In “Transformational Issues of Big Data in Networked Business,” Baesen et al. argue for the addition of a fifth V as a complement to the 4Vs from a business perspective. They suggest that adding the V of ‘value’, will help drive consideration of what to do inside the perimeters set by the 4V definition, namely, how to innovatively investigate and analyze big data and how to anticipate and leverage the transformational impacts of big data.

It would not be hyperbolic to claim that big data is possibly the most significant ‘tech’ disruption in business and academic ecosystems since the meteoric rise of the Internet and the digital economy. With most of today’s data coming from beyond the corporate boundaries, being generated in unstructured format from networks of people and devices, we see more powerful algorithms and better knowledge representation schemes for making sense of all of this heterogeneous and fragmented information than ever before. Since the big data being generated today comes from mostly unstructured sources and networks of people and devices, there is a need for even more powerful algorithms in order to process and make sense of this information.

The definition of information management is constantly evolving as technology, ideas, and business needs change. Data has a lifecycle that is based on its level of usefulness to an organization. Data has to be interpreted to render information, and information has to be understood to emerge as knowledge, which leads to effective decision-making. Big data adds a layer of complexity to the management of information, yet at the same time adds opportunities to derive actionable insights never before possible. Today’s information management practitioners can prosper by taking advantage of the new opportunities made possible by this wealth of data, particularly in leveraging insights.
from data to help organizations make better decisions and create more value for both the customer and for the organization itself.

As the fuel that drives Internet of Things (IoT) and artificial intelligence, big data and analytics can predict problems before they happen. Together with IoT, big data predictive analytics can assist organizations to be proactive instead of reactive. This is incredibly valuable to organizations and can change the way decisions are made which has a cascading effect on all areas of an organization. Given the nature of the profession, information management practitioners are perfectly positioned to help their organizations tackle the challenges of the big data era. With this era comes greater risk, particularly in the area of cybersecurity. Since using analytical techniques on big data creates even more value from information than ever before, information will become one of the greatest commodities that an organization has. Information management practitioners need to expand their security and privacy knowledge, increase their computer literacy and increase their knowledge of risk management in order to better serve their organizations and help to protect the organization’s data and information.

**Organizational and Societal Impact of Big Data**

Making analytical inferences from data is as old as the field of statistics dating back to the 18th century, but today’s inferences are unique since economic and social transactions are quickly moving online, allowing for the digital capture of big data that draws insights from a pool of data sets on human discourse. From the perspective of scientific inquiry, the entire globally connected networked economy now can be envisaged as a large-scale real-world laboratory where researchers can design and conduct experiments and collect the data needed to obtain answers to a variety of questions.

With the incorporation of big data analytical techniques, more organizational decisions will be automated which further impacts the decision processes and responsibilities throughout organizations. Enhanced decision-making is a positive outcome of big data for business as big data allows us to leverage both prediction and causal analysis and draws on a variety of techniques from machine learning, classical statistics, and econometrics, to the design of experiments that existing theories and hypothesis. By
testing these theories, analytics applications have already started to make a measurable impact on organizations\textsuperscript{11}.

Beyond enhanced decision-making in commercial settings, there is great societal benefit from such experimentation since we now have unprecedented digitization of social processes. Several recent editorials have expanded on the idea of the societal impact of big data and analytics. In their article “Big Data & Analytics for Societal Impact: Recent Research and Trends,” Gupta et al. highlight the ability of the innovative uses of big data to have a measurable societal impact, citing in particular the applications of online-to-offline commerce, proactive customer care, and Internet of Things enabled cars\textsuperscript{12}.

While the effects are often positive, there are challenges inherent in the manipulation of data for experimentation which have raised controversies particularly in the social media sphere as in 2014, when Facebook and OKCupid were questioned for experimenting on unaware users by adjusting the content presented to them\textsuperscript{13}. The key question coming out of these controversies is whether such social media experimentation is beneficial to society at large and what moral scrutiny is appropriate in this situation. Current literature indicates that there is a strong case in favour of such experimentation for businesses in particular, not just to avoid costly bad decisions, but also for the pursuit of a better understanding of what drives human social interactions\textsuperscript{14}. For example, manipulating social media data for experimentation can enhance a company’s decision-making and allows for testing social science theories that would otherwise be too costly to test using other means\textsuperscript{15}. While it may benefit the business, the ethical and privacy implications of such experimentation is of concern, particularly when it comes to personal privacy which will be discussed later in this article.

**Decision-Making and Trust**

Big data is more increasingly becoming a trending practice that organizations use to improve and enhance decision-making, thus increasing operational efficiencies and gaining competitive advantage over rivals. The key advantage of using big data and analytics in an organizational context is making better data-driven decisions by linking a
set of explanatory variables to a business response or outcome and thus facilitating predictive or causal inference-based decision-making\textsuperscript{16}. Trust is recognized as one of the most important factors in addition to information (data) which has traditionally garnered the most attention in the literature. The concept of trust in big data and analytics is a key factor in implementing decision-linked inferences from data. Even if the data itself is of the highest quality, an organization will fail to enhance its decision-making if the firm’s leaders do not trust the data and choose not to rely upon either the data or the relevant analytic techniques\textsuperscript{17}. A challenge related to big data and analytics is generating trust in an analytical model, since the performance of a model is summarized using statistical measures that may be difficult for end-users or non-experts to understand. Given that the use of analytical models plays a key role in the decision-making of the firm, it is important that the models are understandable to decision-makers in the organization\textsuperscript{18}.

This challenge has been further explored in the literature and in business, for example, in 2016, KPMG International commissioned Forrester Consulting to examine the power of trust in data and analytics and released the report “Building Trust in Analytics: Breaking the Cycle of Mistrust in D&A.” The report reveals that just 38 percent of organizations have a high level of confidence and trust in their analytics\textsuperscript{19}. As the report summarizes, given the power that it holds, trust in data analytics should be a non-negotiable business priority. Additionally, the report summarizes the ways that KPMG believes organizations must think about trusted analytics as a strategic way to bridge the gap between decision-makers, data scientists and customers to deliver sustainable business results\textsuperscript{20}. Further to generating trust, new challenges compounded by the data revolution require a shift in organizational culture to adapt to the new world of data since it is now possible to experiment cheaply and accurately and base decisions on that data\textsuperscript{21}.

While moving in quietly, technological change has a relatively disruptive long-term impact on business. The shift from the use of in-house database management systems to cloud-based options means that corporate boundaries are stretching. The big data era allows organizations to leverage the power of prediction and causal analysis in order
to develop the corporate body of knowledge. With the many collaboration tools and sharing methods that exist today, businesses have more opportunities to share knowledge and insights with others. In order to be successful, business intelligence and analytics projects depend on significant business or domain knowledge as well as effective communication skills. Long-lasting benefits from such projects are realized when the meaningful and actionable knowledge gained from turning data into information (using analytics) is effectively communicated to the business and domain experts in the organization. This sharing of knowledge is critical to informing decisions within organizations. Knowledge sharing both inside and outside the corporate boundaries is an effective means of gaining competitive advantage in today’s digitalized world.

**Disruptive Impacts of Big Data**

Historically, the Information Systems (IS) field has developed research at the intersection of computing technology and data in business and society and IS researchers focused on problems and outcomes from a broad perspective of the enterprise as a whole. From sociology to political science, to economics and psychology, the IS discipline has been thinking and researching questions at the intersection of technology, data, business and society for five decades. The aforementioned cross-disciplinary nature of IS research is a strength in today’s connected world and positions researchers well to exploit the big data opportunity. With access to even richer data sets and larger volumes of data, there is more potential than ever before for IS thought leadership to become foundational in education, business and policy.

New disruptive technologies such as data analytics come with many opportunities. Dramatic shifts are happening across many professions as job categories will become obsolete in favour of new and emerging jobs created by big data innovations. Data analytics offers a competitive advantage by analyzing an organization’s data quickly and effectively for future insights. Due to this shift, there is a need to retrain employees in this big data age, since there is a training gap in terms of what they need to know and the training they have had up to this point. Big data has disrupted old business methods
and brings the need to develop new employee skillsets including knowledge of cybersecurity, privacy, data protection and risk management.

Even before the development of the cloud and the concept of blockchain technology which has had pervasive impacts in more recent years, IT architectural innovation had a radical impact on development processes and their outcomes and systems developers had to navigate the difficulties of integrating past standalone systems into the world of the Internet\textsuperscript{24}. As organizations are currently in the process of navigating the world of big data, machine-learning skills are becoming more and more necessary for IT professionals in order to build decision systems that are automated. Information management professionals need to be aware of the basic concepts of data analytics and machine learning, particularly as it pertains to the use of information insights that come from big data innovations. Information professionals can create value for their organizations by assisting in the organization, storage and preservation of the outputs of machine learning knowledge for effective decision-making and use in organizations. Communication and knowledge sharing is important to allow for productiveness in this era of radical innovation.

Data analytics has had a dramatic impact on the field of social sciences as collaboration between researchers, social scientists and data scientists makes innovative projects possible. Analysis of micro-level big data offers an opportunity for social scientists to develop more complex psychological models with useful implications of behaviour\textsuperscript{25}. Jennifer Golbeck’s work is a good example of the integration of social sciences and data sciences. In “Predicting Personality from Social Media Text,” Goldbeck uses social media text to predict personalities using psycholinguistic text analysis. Replicating the text-based Big Five personality score predictions generated by Receptiviti API, she uses a prediction algorithm with social media data sets of personality scores for about 9000 users to determine the accuracy of the Receptiviti API Big Five personality predictions\textsuperscript{26}. The outcome of this research was that remarkably similar results were obtained from the four data sets that Goldbeck tested using the prediction algorithm.
The potential for big data innovations using data from social websites can enable researchers to go where they could only dream of going in the past since the vast amount of micro-level big data about human interactions offers opportunities never available in the physical world due to cost or infeasibility of comparable data collection. Social media big data and analytics has been examined more recently in the research regarding trends in human resource management and recruitment. The use of machine learning and predictive analytics, combined with social media big data is changing the way recruitment is carried out in the organizational context. In the field of recruitment, big data is a potential solution to disintermediating talent requirement and availability to enable a better understanding of candidate’s behaviours and expected fit within an organization. The use of social media platforms to link various data streams using appropriately defined unique identifiers is key in using big data to determine candidate behaviours and expected fit in an organization.

It is important to note that informational usefulness is directly proportional to the quality of the data. Far too often companies and universities consider investments in data quality as too expensive and/or too difficult. The importance of having accurate data to maximize the impact of big data techniques cannot be overstated. When it comes to business priorities, data quality is too often relegated to a lower priority in an organization. In the age of big data, this can have catastrophic consequences. Therefore, management should be on board to help enforce data quality processes and procedures throughout an organization. With the wealth of industry-specific knowledge that can be extracted from high quality data, combined with the use of relevant analytic techniques, an investment in data quality can be expected to bring returns well into the future.

**Big Data Security and Privacy Challenges**

More and more organizations need to be aware of the privacy implications and security threats of managing systems in the era of big data. Many of the security applications designed for smaller volumes of data cannot manage the larger volumes of data that are characteristic of the big data age. Coupled with this, the increasing use of cloud-based storage solutions has facilitated more data mining and collection of data than in the past.
With the prevalence of the Internet and cyberattacks in recent years, cybersecurity is an increasingly important area of practice since most people are not aware of how important it is to protect their information on the Internet. Particularly in the organizational context, cybersecurity awareness training is an extremely important tool to ensuring the security of an organization’s information assets. When it comes to cybersecurity awareness, not all threats are created equal. Phishing is a form of cyberattack that is particularly common and happens through email. In the case of phishing, employees are often given training to help them to identify and report suspicious emails. A key to combatting phishing attacks is educating individuals of the signs of a malicious email through cybersecurity awareness training efforts which emphasize this particular form of attack.

In the case of big data privacy, the raw material or data used to glean interesting insights is increasingly heterogeneous and unstructured. With the proliferation of smart technology, an individual’s claim to personal privacy is virtually nonexistent. People largely supply the data used for big data and analytics through their everyday actions. The challenge in big data and privacy is how to protect the privacy of individuals without a significant loss in informational quality. In their paper, “Privacy and Big Data: Scalable Approaches to Sanitize Large Transactional Databases for Sharing,” Menon and Sarkar present a scalable approach to sanitize transactional data. They show that using this heuristic approach allowed sensitive item sets to be removed, while maintaining similar recommendation accuracy as the original rate observed. Such approaches have the potential to be expanded and transform the culture of hesitation around data sharing due to privacy implications like the unintended exposure of sensitive information.

**Transformational Impacts and Information Management Opportunities**

This research article focuses on the transformative power of big data and analytics in today’s business environment and highlights the challenges the area is bringing to light. It explores the technical and managerial issues of business transformation and disruption that require further exploration in this new age of intelligent information. Further, by closely analyzing the more recent big data and analytics research, we examined the organizational and societal impacts of big data, the importance of
decision-making and trust, the disruptive impacts of big data and big data security and privacy challenges. Based on our research, we conclude that the transformative power of big data and analytics will only continue to progress and grow as the field continues to transform business and society in ways that are yet to be realized.

For the information management profession, big data and analytics means information management practitioners need to consider adding new skills to their current skillsets in order to assist organizations with the new security, privacy, risk management and data management issues that are characteristic of the big data age. In this new business landscape, the protection of information and data becomes crucial to an organization. Therefore, the information management strategy needs to be redefined to address the emerging data dynamic that impacts the storage and use of information, which will include greater focus on the protection of information from unauthorized access and use. Given larger volumes of data, information management professionals need to familiarize themselves with options for cloud storage as well as privacy and security issues relevant to cloud storage. Further, cybersecurity becomes of utmost importance as cyberattacks are increasing and will only continue to increase in the big data age where an organization’s data has more value than ever before. For this reason, information management practitioners will need to work closer with privacy, IT, security and legal in their organizations.

Information is power, and this statement has never been more accurate than in this current big data age with data flowing from audio, video, sensors, social media, machines and other sources. All of this data needs to be categorized, stored and protected according to its value and that is a key area where the profession of information management is perfectly positioned to help drive big data towards actionable intelligence, and subsequently provide tremendous value to organizations.
Bibliography


Works Cited


**Biographies**

*Christy Walters*, CIP has over ten years of progressive experience providing information management (IM) advisory and consulting services to clients in several industries. She has a range of consulting experience working with provincial and federal Governments, in the private sector for Deloitte and PwC and, presently, in a University setting while also completing her MBA part-time. Christy is the current Region Director of ARMA Canada Region.

*Crystal Walters* is a Programmer Consultant with many years of experience in data management and electronic information management in the public sector, manufacturing industry, and in higher education. She is focused on strategic ways to manage and analyze data and information for relationship management and philanthropic success. She holds a BA, a post-diploma in Information Management, and is also a business graduate student at Memorial University.