RAE Journal of Business Management

RESEARCH AND KNOWLEDGE
V. 59, N. 6,
November–December 2019

fgv.br/rae

FGV | SAO PAULO SCHOOL OF BUSINESS ADMINISTRATION

FORUM

The journey has just begun
William Lekse

Beyond technology: Management challenges in the Big Data era
Eduardo de Rezende Francisco | José Luiz Kugler | Soong Moon Kang | Ricardo Silva | Peter Alexander Whigham

Information management capability and Big Data strategy implementation
Antonio Carlos Gastaud Maçada | Rafael Alfonso Brinkhues | José Carlos da Silva Freitas Junior

Intention to adopt big data in supply chain management: A Brazilian perspective
Maciel M. Queiroz | Susana Carla Farias Pereira

Measuring accessibility: A Big Data perspective on Uber service waiting times
André Insardi | Rodolfo Oliveira Lorenzo

Factors affecting the adoption of Big Data analytics in companies
Juan-Pedro Cabrera-Sánchez | Ángel F. Villarejo-Ramos

PERSPECTIVES

Big Data and disruptions in business models
Eric van Heck

Plus ça change, plus c'est la même chose
Flavio Bartmann

ESSAY

Corporate crimes: the specter of genocide haunts the world
Cintia Rodrigues de Oliveira

CONTENTS

EDITORIAL

372 DATA AND OPEN SCIENCE
Propriedade dos dados e ciência aberta
Propiedad de datos y ciencia abierta
Maria José Tonelli | Felipe Zambaldi

FORUM / *FÓRUM* / *FORO*

374 THE JOURNEY HAS JUST BEGUN
A jornada acaba de começar
El viaje acaba de empezar
William Lekse

375 BEYOND TECHNOLOGY: MANAGEMENT CHALLENGES IN THE BIG DATA ERA
Além da tecnologia: Desafios gerenciais na era do Big Data
Más allá de la tecnología: Desafíos de gestión en la era de Big Data
Eduardo de Rezende Francisco | José Luiz Kugler | Soong Moon Kang | Ricardo Silva | Peter Alexander Whigham

379 INFORMATION MANAGEMENT CAPABILITY AND BIG DATA STRATEGY IMPLEMENTATION
Capacidade de gestão da informação e implementação de estratégia de Big Data
Capacidad de gestión de la información e implementación de estrategia de Big Data
Antonio Carlos Gastaud Maçada | Rafael Alfonso Brinkhues | José Carlos da Silva Freitas Junior

389 INTENTION TO ADOPT BIG DATA IN SUPPLY CHAIN MANAGEMENT: A BRAZILIAN PERSPECTIVE
Intenção de adoção de big data na cadeia de suprimentos: Uma perspectiva brasileira
Intención de adopción de big data en la cadena de suministros: Una perspectiva brasileña
Maciel M. Queiroz | Susana Carla Farias Pereira

402 MEASURING ACCESSIBILITY: A BIG DATA PERSPECTIVE ON UBER SERVICE WAITING TIMES
Medindo a acessibilidade: Uma perspectiva de Big Data sobre os tempos de espera do serviço da Uber
Medición de accesibilidad: Una perspectiva de Big Data sobre los tiempos de espera del servicio de la Uber
André Insardi | Rodolfo Oliveira Lorenzo

415 FACTORS AFFECTING THE ADOPTION OF BIG DATA ANALYTICS IN COMPANIES
Fatores que afetam a adoção de análises de Big Data em empresas
Factores que afectan a la adopción del análisis de Big Data en las empresas
Juan-Pedro Cabrera-Sánchez | Ángel F. Villarejo-Ramos

PERSPECTIVES / *PERSPECTIVAS*

430 BIG DATA AND DISRUPTIONS IN BUSINESS MODELS
Big Data e disrupções nos modelos de negócios
Big Data y disrupciones en los modelos de negocio
Eric van Heck

433 PLUS ÇA CHANGE, PLUS C'EST LA MÊME CHOSE
Quanto mais as coisas mudam, mais elas permanecem as mesmas
Cuanto más cambian las cosas, más permanecen igual
Flavio Bartmann

ESSAY / *PENSATA* / *ENSAYO*

435 CORPORATE CRIMES: THE SPECTER OF GENOCIDE HAUNTS THE WORLD
Crimes corporativos: O espectro do genocídio ronda o mundo
Crimen corporativos: El espectro del genocídio alrededor del mundo
Cintia Rodrigues de Oliveira

CC BY

ISSN 0034-7590; eISSN 2178-938X

EDITORIAL

Translated version
DOI: http://dx.doi.org/10.1590/S0034-759020190601



Maria José Tonelli
Editora-chefe



Felipe Zambaldi
Editor-adjunto

DATA AND OPEN SCIENCE

The practice of providing open access to articles, adopted in Brazil and several other countries, still faces resistance from many commercial publishers abroad (Packer & Santos, 2019a). Recently, some of them have switched to a hybrid model (open and closed access) as a more balanced path. Transparency in the peer review process is also being discussed: The digital library SciELO recommends a "gradual increase of transparency and openness [...] with the disclosure of the identities of authors and reviewers during the evaluation process" (Packer & Santos, 2019b). Even more controversial is the policy of open access to research data of articles published in scientific publications. The *Blog SciELO em Perspectiva* has several texts discussing this trend, which have been questioned by many actors involved in the production and publication of scientific articles, including authors, universities, editors, and publishers. Who owns the data? Nassi-Calò (2019) reveals that research on this issue remains inconclusive, and many actors in this process may be the owners, such as funders of the study, institutions of the researcher, publishers and, of course, the authors of the study. The author argues that open science "is demanded by society, governments, and sponsors. This practice brings several advantages by making science more transparent, reproducible, reliable, and verifiable" (Nassi-Calò, 2019). However, several questions arise regarding researchers. In qualitative research conducted by means of interviews, for example, when the anonymity of the interviewees is ensured by an informed consent form, how does one proceed? Although this may not apply to the Exact and Biological Sciences, it is urgent for research in the Human Sciences, because participants could be identified, thus violating the confidentiality ensured within the ethical standards of the study. Besides that, it's also consider the necessary time and resources of the researchers, as well as the ownership of secondary data from third parties that often only give access to only one specific study. Other aspects related to data transparency in this open science era, described as e-science, include the need for cyber-structure (technological bases that support the data), the collaboration of society, as well as the support of the State, as expressed by Targino and Garcia (2018). But again, who owns the technology infrastructure that stores the data? Packer and Santos (2019b) argue that open science is an irreversible movement, and the 4th Brazilian Action Plan on this topic involves some clearly defined milestones for the future based on the guidelines of the Global Open Fair. Although these guidelines have already been implemented in the field of healthin Brazil, the authors argue that graduate programs should invest in training programs. The State of São Paulo Research Foundation (*Fundação de Amparo à Pesquisa do Estado de São Paulo* [FAPESP]) considers this orientation in Thematic Projects. In the future, the quality of articles will be assessed not only by the journal in which they are published but also the available data (Kiley & Markie, 2019). This action is expected to eliminate problems such as plagiarism, reproducibility of the study, and biases. These criteria are undoubtedly valid for the Exact and Biological Sciences, but are they applicable to the Human Sciences, which are often concerned with unique and non-replicable phenomena? If the neutrality of algorithms is questioned even today, can we really engage in science without bias? Are data neutral?

 CC BY ISSN 0034-7590; eISSN 2178-938X

This edition includes a forum on Big Data, organized by Eduardo de Rezende Francisco, José Luiz Kugler, Soong Moon Kang, Ricardo Silva, and Peter Alexander Whigham. The first guest article presented is "The journey has just begun" by William Lekse. Following the introduction to the forum, the next article presented is "Beyond technology: Managing challenges in the Big Data era" by the guest editors. The articles presented after this include: "Information management capability and Big Data strategy implementation" by Antonio Carlos Gastaud Maçada, Rafael Alfonso Brinkhues, and José Carlos da Silva Freitas Junior; "Intention to adopt big data in supply chain management: A Brazilian perspective" by Maciel M. Queiroz and Susana Carla Farias Pereira; "Measuring accessibility: A Big Data perspective on Uber service waiting times" by André Insardi and Rodolfo Oliveira Lorenzo; and "Factors affecting the adoption of Big Data analytics in companies" by Juan-Pedro Cabrera-Sánchez and Ángel F. Villarejo-Ramos. The Perspectives section raises the debate on the use of Big Data in business through articles such as "Big Data and disruptions in business models" by Eric Van Heck, and "*Plus ça change, plus c'est la même chose* [The more things change, the more they remain the same]" by Flávio Bartman. The essay "Corporate crimes: The specter of genocide haunts the world" by Cintia Rodrigues de Oliveira reminds us that misconduct, unethical behavior, and corporate social irresponsibility also permeate the business world.

Happy reading!

Maria José Tonelli[1] | ORCID: 0000-0002-6585-1493
Felipe Zambaldi[1] | ORCID: 0000-0002-5378-6444
[1]Fundação Getulio Vargas, São Paulo School of Business Administration, São Paulo, SP, Brazil

REFERENCES

KILEY, R., & MARKIE, M. (2019). Wellcome Open Research, o futuro da Comunicação Científica? [Publicado originalmente no blog LSE Impact of Social Sciences em fevereiro/2019] [online]. SciELO em Perspectiva. Retrieved from: https://blog.scielo.org/blog/2019/02/27/wellcome-open-research-o-futuro-da-comunicacao-cientifica/

Nassi-Calò, L. (2019). Promovendo e acelerando o compartilhamento de dados de pesquisa [on-line]. SciELO em Perspectiva. Retrieved from https://blog.scielo.org/blog/2019/06/13/promovendo-e-acelerando-o-compartilhamento-de-dados-de-pesquisa/

Packer, A. L., & Santos, S. (2019a). Ciência aberta e o novo modus operandi de comunicar pesquisa – Parte I [on-line]. SciELO em Perspectiva. Retrieved from https://blog.scielo.org/blog/2019/08/01/ciencia-aberta-e-o-novo-modus-operandi-de-comunicar-pesquisa-parte-i/

Packer, A. L., & Santos, S. (2019b). Ciência aberta e o novo modus operandi de comunicar pesquisa – Parte II [on-line]. SciELO em Perspectiva. Retrieved from https://blog.scielo.org/blog/2019/08/01/ciencia-aberta-e-o-novo-modus-operandi-de-comunicar-pesquisa-parte-ii/

Targino, M. G., & Garcia, J. C. R. (2018). Perspectivas da avaliação por pares aberta: Instigante ponto de interrogação [on-line]. SciELO em Perspectiva. Retrieved from https://blog.scielo.org/blog/2018/05/14/perspectivas-da-avaliacao-por-pares-aberta-instigante-ponto-de-interrogacao/

 CC BY ISSN 0034-7590; eISSN 2178-938X

FORUM

Invited article
Original version
DOI: http://dx.doi.org/10.1590/S0034-759020190602

THE JOURNEY HAS JUST BEGUN

This special issue contributes to incorporating research utilizing Big Data —in particular, the disciplines of information systems and supply chain—into mainstream academic research. It also extends Big data's contribution to establishing predictive analytics-based research as theory building. Big Data brings many features to academic research that, if properly understood, can shift the approach of most research towards that of the classical academic approach, which focuses on building and testing theory. The academic approach to theoretical research seeks to explain phenomena by applying frameworks, which are sourced from different disciplines such as microeconomics, operations research, organizational theory, psychology, and sociology.

Of particular interest to academic researchers and practitioners alike is the capability to analyze, explain, and predict consumer behavior. Most retail sales still occur in stores, and consumers who purchase online also visit stores before or after a sale. Presently, the majority of store shoppers use mobile devices to perform research on products, communicate with family and friends, and visit sites, which provide data - often, Big Data - to facilitate the shopping experience (Fildes & Kolassa, 2018). Thus, not only transactions or lack thereof, but also all technological aspects of the shopping experience, can now be extensively modeled. Much of this data is now starting to become available to academic researchers. Therefore, technology, particularly data collection, processing, and dissemination of Big Data, is making significant contributions in the global marketplace.

These new technologies and trends in Big Data are emerging in local, regional, and global consumer behavior analysis and extending throughout supply chain operations. Big Data is changing the rules of business—from the design and prototyping to the production and distribution of products and services. Academics now have what they have long required: sources of massive quantities of data. For decades, consumer behavior investigations in journal publications were limited to researcher-generated datasets. The larger Big Data datasets offer academic researchers and practitioners the means to become more aware of relevant updates on a real-time basis. Researchers no longer need time to develop a research plan - which includes specifying frameworks, developing models, and seeking permissions - to perform investigations on small samples. Researchers can now investigate multiple models and frameworks of theories integrated from several disciplines. The means to test explanatory as well as predictive theories (including grounded theory) is now available to researchers around the globe. Moreover, every investigation can be rapidly replicated and verified as the data is available to all academics, allowing a more productive and creative global research environment (Johnson, Gray & Sarker, 2019).

This special issue explores different methods to tackle on relevant analytical challenges. This exciting journey has just begun, and will certainly lead to interesting research avenues.

WILLIAM LEKSE[1]
wjlekse@katz.pitt.edu
ORCID: 0000-0002-9972-3393

[1]University of Pittsburgh, Joseph M. Katz Graduate School of Business & College of Business Administration, Pittsburgh, PA, United States of America

REFERENCES

Fildes, R. A., & Ma, S., Kolassa, S. (2018). Retail forecasting: Research and practice. *Management Science*. Working paper. Unspecified, Lancaster, UK.

Johnson, S. L., Gray, P, & Sarker, S. (2019), Revisiting IS research practice in the era of big data, Information & Organization, *29*(1), 41-56. doi:10.1016/j.infoandorg.2019.01.001

 CC BY ISSN 0034-7590; eISSN 2178-938X

FORUM

Invited article
Original version
DOI: http://dx.doi.org/10.1590/S0034-759020190603

BEYOND TECHNOLOGY: MANAGEMENT CHALLENGES IN THE BIG DATA ERA

EDUARDO DE REZENDE FRANCISCO[1]
eduardo.francisco@fgv.br
ORCID: 0000-0001-8895-2089

JOSÉ LUIZ KUGLER[1]
jose.kugler@fgv.br
ORCID: 0000-0003-1625-7807

SOONG MOON KANG[2]
smkang@ucl.ac.uk
ORCID: 0000-0003-1605-601X

RICARDO SILVA[3]
ricardo.silva@ucl.ac.uk
ORCID: 0000-0002-6502-9563

PETER ALEXANDER WHIGHAM[4]
peter.whigham@otago.ac.nz
ORCID: 0000-0002-8221-6248

[1]Fundação Getulio Vargas, Escola de Administração de Empresas de São Paulo, São Paulo, SP, Brazil

[2]University College London, School of Management, London, United Kingdom

[3]University College London, Department of Statistical Science, London, United Kingdom

[4]University of Otago, Department of Information Science, Dunedin, Otago, New Zealand

INTRODUCTION

The ability of organizations to produce, collect, manage, analyze, and transform data has increased rapidly over the past decade (Delen & Zolbanin, 2018). This has resulted in significant new challenges regarding how data can be leveraged for improving business decisions and how this new scenario changes business processes and operations (Vidgen, Shaw, & Grant, 2017). The widespread adoption of advanced analytical methods (e.g., machine learning) has attracted significant interest (Gupta, Deokar, Iyer, Sharda, & Schrader, 2018; Vassakis, Petrakis, & Kopanakis, 2018) particularly because the required data storage and methods can be accessed remotely through web-based interfaces such as cloud services. This has resulted in an increased belief that businesses must actively engage with this technology to remain competitive. However, this Red Queen scenario comes at a cost as collecting, curating, and managing large datasets requires expertise and dedicated staff, often consuming resources that do not contribute to core business activities. Consider the fact that there is an increasing role for data scientists and data engineers, among others, within organizations (Davenport & Patil, 2012). Roles such as Chief Data Officer (CDO) and Chief Analytics Officer (CAO) are now commonplace within most organizations.

There is also the issue regarding data preparation. The mantra that 80% of the effort is in the management of data is still largely correct. In addition, the appropriate use and interpretation of predictive models requires expertise that involves both a deep understanding of the underlying business and the assumptions and limitations of each model. Finding suitable people that have skills from both a business and technology perspective can be difficult. The cost-benefit trade-off for businesses in relation to Big Data is often difficult to assess and may lead to failures in how the proposed solution is developed and linked to the business model (Loebbecke & Picot, 2015). There are many examples of organizations, especially government and publicly funded organizations, in which scarce resources may be wasted on failed analytical projects because of a misunderstanding about how the data is meant to be used, the types of data that are collected, and the questions the model intends to address. There are also relevant issues with slow development lifecycles in the analytical arena. Since technology is evolving at a rapid rate, a project that takes a significant amount of time may result in a solution that is expensive compared with a solution using the latest technology. Knowing when to develop, and when to wait, is also a key challenge to the current mechanisms of analytical governance.

 CC BY ISSN 0034-7590; eISSN 2178-938X

There is a need to understand how organizations should transform their business models when confronted with this increasingly rich world of data, and how they can ensure compliance with correct practices not only from the perspective of technology but also from the managerial, ethical, and societal viewpoints. Early discussions on the theme of Big Data were often framed around the V's perspective (volume, velocity, variety, value, veracity, variability, visualization) (Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015), and although these concepts still remain relevant today there is increasing acknowledgement that data is not a disconnected concept. This has led to the notion of managing data from an ecosystem perspective (Demchenko, de Laat, & Membrey, 2014). Broadly, a natural ecosystem operates at a range of spatial and temporal scales from the individual within a species, to the species community, to food webs and the environment, all within the context of both exogenous drivers (such as climate and competition) and endogenous factors (such as nutrient requirements). Data can also be seen within this broader framework (Gupta et al., 2018) and should therefore be used and modelled as part of a larger, dynamic system, rather than as a separate, disconnected concept. This includes the origin from which data is collected, other digital devices and sensors, technology providers and broader communities involved in data creation, policy making, and so on.

What are the current trends in Big Data analytics? There are two main directions worth mentioning in relation to business decision making: Integrated data infrastructures (IDIs) and the internet of things (IoT) (Ahmed et al., 2017). The main concept of IDIs is that linking or associating data together may provide additional opportunities for examining the structure and relationships between all of these datasets. Many government organizations have collected data through separate organizations, such as justice, health, education, income, social services, community, and population statistics (e.g., regular census collections). However, until recently, most of these data could not be linked in a useful way and it was difficult to obtain a common format or gain access to these types of data. IDIs allow these types of data to be used together, allowing an ecosystem view of society to emerge by presenting person-centric microdata that can be related to aggregated data. Understanding how individuals interact, how decisions are made by individuals, and how these are reflected in societal outcomes (Newell & Marabelli, 2015) allows a greater understanding of why people behave under different circumstances. This means that businesses must understand the way in which social structures, from the individual to the societal perspective, are operating, and therefore the business opportunities that can be leveraged from the individual perspective. IDIs allow questions to be addressed in areas as diverse as agricultural production, mental health, education development, the labor market, immigration, tourism, wage disparities, gender inequality, and so on. However, there are complications with using an IDI; security issues mean that access is often limited or highly controlled, and access for business purposes may be limited unless there is a direct association with a research organization such as a university. However, the current trend in developing IDIs means that businesses will ultimately benefit from these linked data sources, whether it be for their own purposes or as a provider of tools and methods for integrating and using such data sources.

The Internet of Things (IOTs) (Ahmed et al., 2017) continues to be driven by consumer demand with the promise of improved personalization of services and control over many individual decision-making processes. The eventual rollout of fast 5G wireless networks and the increase in connectivity among all devices (remember when the smart-phone was introduced, but now it is just a phone) will lead to business opportunities in terms of how these connections are used, what they represent from an individual perspective, and which new products and services can be created around this ecosystem. Access to this data will also allow new approaches to understanding individual behavior, how consumer demand is created (Erevelles, Fukawa, & Swayne, 2016), and methods for optimizing how individuals interact with systems. Smart electronic devices also allow local processing to be performed; the notion of edge computing and the pre-processing of data to filter and reduce how information is used will become a fundamental aspect of IOT development. Business opportunities exist from both the hardware and software perspective, from what types of devices will be used to how they will interact as a system. The world of sensors and how this will change our perspective in terms of business opportunities has only just begun.

This huge, unprecedented influx of data offers a plethora of opportunities. However, to leverage such opportunities we need to develop meaningful models; to make sense of the complexity that characterizes our economic, political, and social challenges we need to develop sensible, well-articulated models that attempt to reveal how causal processes overlap and interact (Page, 2018).

From a management perspective, the mission is how to recognize which business processes can benefit from what kind of models, how the data can be organized and used, and how analytical results can be incorporated into the decision-making framework.

 CC BY © RAE | São Paulo | **59(6)** | **November-December 2019** | 375-378 ISSN 0034-7590; eISSN 2178-938X

ACCEPTED ARTICLES

In this special issue, we focus mainly on the most basic of these challenges, namely, the decision to implement Big Data technologies. In "Factors affecting the adoption of big data analytics in companies," Cabrera-Sánchez and Villarejo Ramos (2019) examine the barriers to implement Big Data techniques based on online survey of managers in different areas such as marketing, finance, and human resources. They found that companies with little or no experience with Big Data are more prone to social influence, exhibit higher expectations about the new technology and have higher resistance to adopt the new technology, whereas companies with more experience are more interested in easy access and necessary support for the technology and show lower expectations about its performance.

With special attention to experiences in Brazil, in their paper "Intention to adopt Big Data in supply chain management: A Brazilian perspective," Queiroz and Farias (2019) use a similar framework as that employed by Cabrera-Sánchez and Villarejo Ramos (2019), namely the unified theory of acceptance and use of technology (UTAUT), to analyze specifically the intention to adopt Big Data techniques among Brazilian supply chain management professionals who had some experience with the technology. For these professionals, the main factor to adopt the Big Data technology depends on IT infrastructure such as access to high-speed internet and integration with other systems.

In their paper, "Information management capability and Big Data strategy implementation," Maçada, Brinkhues, and Freitas Junior (2019) investigate how an organization's expectations about benefits and costs of Big Data are influenced by its ability to access data and information from its environment, to process them, and to meet the market needs based on them, or "Information Management Capability" (IMC). They demonstrate that IMC is positively related to value expectations and negatively related to cost expectations, which in turn negatively affect the intent to purchase resources and capabilities to implement Big Data.

Finally, as an application of Big Data, Insardi and Lorenzo (2019) in "Measuring accessibility: A Big Data perspective on Uber service waiting times", used some basic Big Data techniques to study mobility access in a large urban setting using estimated waiting times of all Uber products in the city of Sao Paulo. Their major finding is that the estimated waiting times are highly related to socio-economic variables of the neighborhoods (districts). For example, the authors found a strong relationship between the waiting times and the proportion of non-white population.

REFERENCES

Ahmed, E., Yaqoob, I., Hashem, I. A. T., Khan, I., Ahmed, A. I. A., Imran, M., & Vasilakos, A. V. (2017). The role of big data analytics in internet of things. *Computer Networks*, *129*(Part 2), 459-471. doi:10.1016/j.comnet.2017.06.013

Cabrera-Sánchez, J-P., & Ramos, A. F. V. (2019). Factors affecting the adoption of big data analytics in companies. *RAE-Revista de Administração de Empresas*, *59*(6), 415-429. doi: http://dx.doi.org/10.1590/S0034-759020190607

Davenport, T. H., & Patil, D. J. (2012). Data scientist: The sexiest job of the 21st century. *Harvard Business Review*, *90*(10), 70-76

Delen, D., & Zolbanin, H. M. (2018). The analytics paradigm in business research. *Journal of Business Research*, *90*, 186-195. doi:10.1016/j.jbusres.2018.05.013

Demchenko, Y., Laat, C. de, & Membrey, P. (2014). Defining architecture components of the big data ecosystem. *International Conference on Collaboration Technologies and Systems (CTS)* (pp. 104-112). doi:10.1109/CTS.2014.6867550

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897-904. doi: 10.1016/j.jbusres.2015.07.001

Gupta, A., Deokar, A., Iyer, L., Sharda, R., & Schrader, D. (2018). Big data & analytics for societal impact: Recent research and trends. *Information Systems Frontiers*, *20*(2), 185-194. doi: 10.1007/s10796-018-9846-7

Insardi, A., & Lorenzo, R. (2019). Measuring accessibility: A big data perspective on Uber service waiting times. *RAE-Revista de Administração de Empresas*, *59*(6), 402-414. doi: http://dx.doi.org/10.1590/S0034-759020190606

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *Journal of Strategic Information Systems*, *24*(3), 149-157. doi: 10.1016/j.jsis.2015.08.002

Maçada, A. C. G., Brinkhues, R. A., & Freitas, J. C. da S., Junior. (2019). Information management capability and big data strategy implementation. *RAE-Revista de Administração de Empresas*, *59*(6), 379-388. doi: http://dx.doi.org/10.1590/S0034-759020190604

Newell, S., & Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of "datification". *Journal of Strategic Information Systems*, *24*(1), 3-14. doi: 10.1016/j.jsis.2015.02.001

Page, S. E. (2018). *The model thinker: What you need to know to make data work for you*. New York, USA: Hachette Book Group.

Queiroz, M. M., & Farias, S. C. (2019). Intention to adopt big data in supply chain management: A Brazilian perspective. *RAE-Revista de Administração de Empresas*, *59*(6), 389-401. doi: http://dx.doi.org/10.1590/S0034-759020190605

Vassakis, K., Petrakis, E., & Kopanakis, I. (2018). Big data analytics: Applications, prospects and challenges. In G. Skourletopoulos, G. Mastorakis, C. Mavromoustakis, C. Dobre, & E. Pallis (Eds.), *Mobile big data* (Vol. 10, pp. 3-20). doi:10.1007/978-3-319-67925-9_1

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, *261*(2), 626-639. doi:10.1016/j.ejor.2017.02.023

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How "big data" can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, *165*, 234-246. doi:10.1016/j.ijpe.2014.12.031

 CC BY

FORUM

Submitted 10.01.2018. Approved 07.19.2019
Evaluated through a double-blind review process. Guest Scientific Editors: Eduardo de Rezende Francisco, José Luiz Kugler, Soong Moon Kang, Ricardo Silva, and Peter Alexander Whigham
Original version
DOI: http://dx.doi.org/10.1590/S0034-759020190604

INFORMATION MANAGEMENT CAPABILITY AND BIG DATA STRATEGY IMPLEMENTATION

Capacidade de gestão da informação e implementação de estratégia de Big Data

Capacidad de gestión de la información e implementación de estrategia de Big Data

ANTONIO CARLOS GASTAUD MAÇADA[1]
acgmacada@ea.ufrgs.br
ORCID: 0000-0002-8849-0117

RAFAEL ALFONSO BRINKHUES[2]
rafael.brinkhues@viamao.ifrs.edu.br
ORCID: 0000-0002-9367-5829

JOSÉ CARLOS DA SILVA FREITAS JUNIOR[3]
freitas1995@gmail.com
ORCID: 0000-0002-9050-1460

[1]Universidade Federal do Rio Grande do Sul, Escola de Administração, Porto Alegre, RS, Brazil

[2]Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul, Viamão, RS, Brazil

[3]Universidade do Vale do Rio dos Sinos, Escola de Gestão e Negócios, São Leopoldo, RS, Brazil

ABSTRACT

Firms are increasingly interested in developing Big Data strategies. However, the expectation of the value of these benefits and of the costs involved in acquiring or developing these solutions are not homogeneous for all firms, which generates competitive imperfections in the market for strategic resources. Information Management Capability (IMC) aims to provide the required unique insights for successful Big Data strategies. This study analyzes IMC as an imperfection agent in the market for strategic Big Data resources. The hypotheses were tested using a survey of 101 respondents and analyzed with SEM-PLS. The results indicate the positive influence of IMC on value expectation and a negative effect on cost expectation. Cost expectation inversely affects the intent to purchase or develop the resources to implement Big Data strategies. Value expectation has a positive effect on both intents.

KEYWORDS | Big Data, information management, strategic factor market, value expectation, cost expectation.

RESUMO

O interesse das organizações em desenvolver estratégias de Big Data está aumentando significativamente. No entanto, a expectativa do valor desses benefícios e dos custos envolvidos na aquisição ou desenvolvimento dessas soluções não é homogênea para todas as empresas, gerando imperfeições competitivas no mercado de recursos estratégicos. A Capacidade de Gestãoda Informação (CGI) tem como premissa fornecer as informações necessárias para que as estratégias de Big Data sejam bem-sucedidas. Este artigo se propõe a analisar o CGI como um agente imperfeito no Strategic Factor Market de Big Data. As hipóteses foram testadas a partir de uma pesquisa de 101 respondentes e analisadas com a utilização de SEM-PLS. Os resultados indicam uma influência IMC positiva na expectativa de valor e uma negativa na expectativa de custo. A expectativa de custo afeta inversamente a intenção de comprar ou desenvolver os recursos para implantar estratégias de Big Data. A expectativa de valor tem um efeito positivo em ambas as intenções.

PALAVRAS-CHAVE | Big Data, gestão da informação, strategic factor market, expectativa de valor, expectativa de custo.

RESUMEN

El interés de las organizaciones en el desarrollo de estrategias de Big Data está aumentando significativamente. Sin embargo, la expectativa del valor de los beneficios y de los costos implicados en el acreedor o el desarrollo de estas soluciones no es homogénea para todas las empresas, impugnando las imperfecciones en el mercado de los recursos estratégicos. Capacidad de Gestión de la Información (CGI) utiliza las premisas proporcionar las pruebas requeridas para el éxito de Big Data, este artículo tiene como objetivo analizar el CGI como un agente de imperfección en el Strategic Factor Market de Big Data. Las hipótesis se probaron de una encuesta de 101 respondedores y se analizaron con SEM-PLS. Los resultados indican la positiva influencia de CGI sobre la expectativa y una negativa en una expectativa de los costos. La expectativa de los costos inversamente afecta al intento de comprar o de desarrollar los recursos para implementar estrategias Big Data. La expectativa de valor tiene un efecto positivo en ambos intents.

PALABRAS-CLAVES | Big Data, information management, strategic factor market, expectativa de valor, expectativa de los costos.

 CC BY

INTRODUCTION

"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" (Agarwal & Dhar, 2014, p. 443). Diverse forms of data that do not generate value do not contribute to an organization. Data value is, thus, driving increasing interest in big data (Chiang, Grover, Liang, & Zhang, 2018). Researchers and technology vendors recognize the benefits of adopting big data analytics in business practices (Wang, Kung, Wang, & Cegielski, 2018). Firms are increasingly interested in developing Big Data strategies (Tabesh, Mousavindin, & Hasani, 2019). The percentage of firms that already invest or plan to invest in Big Data grew from 64 percent in 2013 (Gartner, 2014) to 73 percent (Davenport & Bean, 2018). "Organizations are currently looking to adopt Big Data technology, but are uncertain of the benefits it may bring to the organization and concerned with the implementation costs" (Lakoju & Serrano, 2017, p. 1). The volume of investments is growing at an even greater rate. The Big Data technology and services market will grow at an 11.9 percent compound annual growth rate (CAGR) to 260 billion dollars through 2022 (IDC, 2018). The expected organizational impacts are many, and include cost reductions, an increase in business insights, revelations of strategic information, and improved decision making (Kwon, Lee, & Shin, 2014). However, the expected value of these benefits and the costs involved to acquire and develop these solutions are not homogeneous for every firm, which generates competitive imperfections in the market for strategic resources.

According to strategic factor market (SFM) theory, firms need to be consistently more informed than are other firms that aim to implement the same strategy to obtain superior performance (Barney, 1986). The author affirms that analyzing the firm's capabilities can create these circumstances more so than the competitive environment. We argue that information management capability (IMC) can bring the unique insight required for successful Big Data strategies. We define IMC as the firm's ability to access data and information from internal and external environments, to map and distribute data for processing, and to allow the firm to adjust to meet the market needs and directions. The literature indicates that IMC positively influences a firm's performance directly (Carmichael, Palácios-Marques, & Gil-Pichuan, 2011) or is mediated by other organizational capabilities (Mithas, Ramasubbu, & Sambamurthy, 2011). There is no evidence that a firm's current IMC can accommodate the sharp growth in the flow of unstructured data (White, 2012).

However, IMC can have a relevant role in the expectations for and intent to implement a strategy to deal with Big Data. Many practitioners are seeking such opportunities due to easy access to computational capabilities and analytical software (Agarwal & Dhar, 2014). On the other hand, 43 percent of directors refer to budget deficits as the main barrier delaying the actions to take advantage of this context (Mckendrick, 2013). This indicates symmetry in the cost expectation of the resources for a Big Data strategy. From an academic standpoint, many studies investigate this phenomenon, especially in Information Systems (IS) in terms of analyzing the value creation from these data (e.g., Brown, Chui, & Manyika, 2011; Davenport, Barth & Bean, 2012; Johnson, 2012; McAfee & Brynjolfsson 2012, Lakoju & Serrano, 2017).

Nevertheless, few works focus on the relationship between IMC and Big Data in order to obtain this value (Brinkhues, Maçada, & Casalinho, 2014; Mohanty, Jagadeesh, & Srivatsa, 2013). "The current literature on big data value realization is characterized by a limited number of empirical studies and some repackaging of old ideas" (Günther, Rezazade Mehrizi, Huysman, & Feldberg, 2017). This study aims to determine how the variation in the level of IMC among the firms creates competitive imperfections in the resources market for the implementation of Big Data strategies. To cover this research gap, we propose a scale to measure IMC and conceptually develop a research model to evaluate the relationship between IMC and the implementation of Big Data strategy empirically. This model, based on SFM theory, specifically investigates the influence of IMC on the value and cost expectations of the resources needed for this implementation, and based on transaction cost theory, the effect of these expectations on the intent to acquire or develop these resources. We constructed the scale following the literature and collect data from executives via card sorting. We tested the research model through a survey of 101 directors and analyze the data utilizing SEM-PLS.

This article proceeds as follows. The next section develops the hypotheses and presents the research model. The following section details the procedures to construct the IMC scale and for data collection. We present and discuss the results thereafter, and finally offer our conclusions and implications for research and managerial practice.

INFORMATION MANAGEMENT CAPABILITY (IMC) AND THE STRATEGIC FACTOR MARKET (SFM)

"Strategic Factor Markets (SFM) are markets where the necessary resources for implementation of a strategy are acquired" (Barney 1986, p. 1231); thus, firms can only extract superior performance when SFM is imperfect due to the differences in the expectation of

the future value of these strategic resources. In other words, firms must be able to exploit a larger value of the necessary resources for its strategic implementation rather than the costs to acquire them being significantly less than their economic value. "The goal of big data programs should be to provide enough value to justify their continuation while exploring new capabilities and insights" (Mithas, Lee, Earley, Murugesan, & Djavanshir, 2013, p. 18). To obtain this advantage, firms need to be consistently better informed than the other firms acting in the same SFM (Barney, 1986). IMC can serve as leverage in this advantage.

Mithas et al. (2011) propose the IMC construct to develop a conceptual model linking it with three other organizational capabilities (customer management, process management, and performance management). Their results show that these management capabilities mediate the positive influence of IMC on the firm's performance. Mithas et al.'s (2011) IMC concept consists of three abilities: to provide data and information to users with appropriate levels of accuracy, timeliness, reliability, security, and confidentiality; to provide connectivity and universal access at an adequate scope and scale; and to adapt the infrastructure to the emerging needs and directions of the market. Carmichael et al. (2011) define IMC as a second-order construct composed of the compilation and production of information; access to information; and the identification of information distribution requirements. Another author, Phadtare (2011), proposes that IMC is linked to five factors: acquisition and retention, processing and synthesis, recovery and use, transmission and dissemination, and support system and integration.

Based on the three works above (Mithas et al., 2011; Phadtare, 2011; Carmichael et al., 2011), we identify five dimensions of IMC (access, distribution, people, architecture, and infrastructure). Then, as we explain in detail in the next sections, we perform a card sorting analysis with executives, which pointed to a 10-item scale of these dimensions. From this analysis, we formulated a definition of IMC and applied in this study as corresponding to the firm's set of skills that articulate information infrastructure, the architecture of information, and access to information, which enable organizational adjustment in response to changes imposed by internal and external environments. Thus, we expect that organizations with more developed IMC are more accurate in their expectations of value and can take advantage of the asymmetry of information in the SFM, from which competitive imperfections in SFM derive.

Additionally, we expect that companies that developed IMC at a higher level during one of the previous eras of IM – Decision Support, Executive Support, Online Analytical Processing, and Business Intelligence and Analytics (Davenport, 2014) – have a higher value expectation of the next frontier of Big Data. We predict this result because the development of IMC at an elevated level positively impacts organizational performance (Carmichael et al., 2011; Mithas et al., 2011), which favors a polarizing effect of perceptions between past and present (Vasconcelos, Mascarenhas, & Vasconcelos, 2006). Big Data strategy is a set of solutions based on recent advances in Big Data analytics. Organizations seek to incorporate these solutions in their own decision-making processes successfully (Tabesh et al., 2019). Hence, these firms have a greater expectation of value from Big Data strategies based on their positive experiences with prior IM investments. Conversely, firms that did not reach the same level of IMC may not have had the same success in their ventures in IM, and this negative experience may reflect in a greater expectation of the cost to adopt this type of strategy.

> H1: Firms with more elevated IMC have a lower cost expectation to implement a Big Data strategy.
>
> H2: Firms with more elevated IMC have greater expectations of value extraction from implementing a Big Data strategy.

Asymmetric value expectation and intent to purchase/develop Big Data strategy capabilities

Prior studies also demonstrate the positive effect of using data for the purpose of acquiring Big Data solutions (Kwon et al., 2014). However, firms can also develop the resources and capabilities to implement a Big Data strategy internally.

Organizations exist to realize internal transactions more efficiently than it is to do so in the market (Coase, 1937). Accordingly, firms that do not arrange their resources to reach their objectives more efficiently than the market lose their reason to exist. Thereby, the search for the necessary resources to implement a Big Data strategy can go down two paths: to develop them internally or to acquire them in the market. Organizations can develop the necessary capabilities internally for this implementation if they are efficient in rearranging the resources involved. However, if the cost to acquire such funds in the market is less than the value to produce them internally, then firms tend to acquire them.

Transactions costs are the consequence of the asymmetrical and incomplete distribution of information among the organizations involved in the exchange (Cordella, 2006). The emergence of various suppliers with solutions to manage Big Data leaves uncertainty about what value firms can exploit from these resources. Thus, the decision to buy or develop the factors necessary to implement a Big Data strategy is also affected by the differences in the asymmetrical expectations of value that the firm can extract from this investment. We expect that different levels of expectations positively influence

 CC BY © RAE | São Paulo | **59(6)** | **November-December 2019** | 379-388 ISSN 0034-7590; eISSN 2178-938X

both decisions, whether to purchase or internally develop the resources to extract value from Big Data.

> H3a: Firms with greater value extraction expectations of Big Data strategies have a higher purchase intent for these solutions.

> H3b: Firms with greater value extraction expectations of Big Data have a higher intent to develop these solutions internally.

Asymmetric cost expectation and intent to purchase or develop Big Data strategy capabilities

Resources such as million instructions per second (MIPS) and terabytes of storage for structured data are less expensive through Big Data technologies than through traditional technologies (Davenport, 2014). However, the costs of other less tangible resources may be more difficult to predict.

For instance, transaction costs frequently increase when adopting an IS solution. However, firms can reduce these costs when the costs associated with adoption do not exceed the external costs that affect adoption (Cordella, 2006).

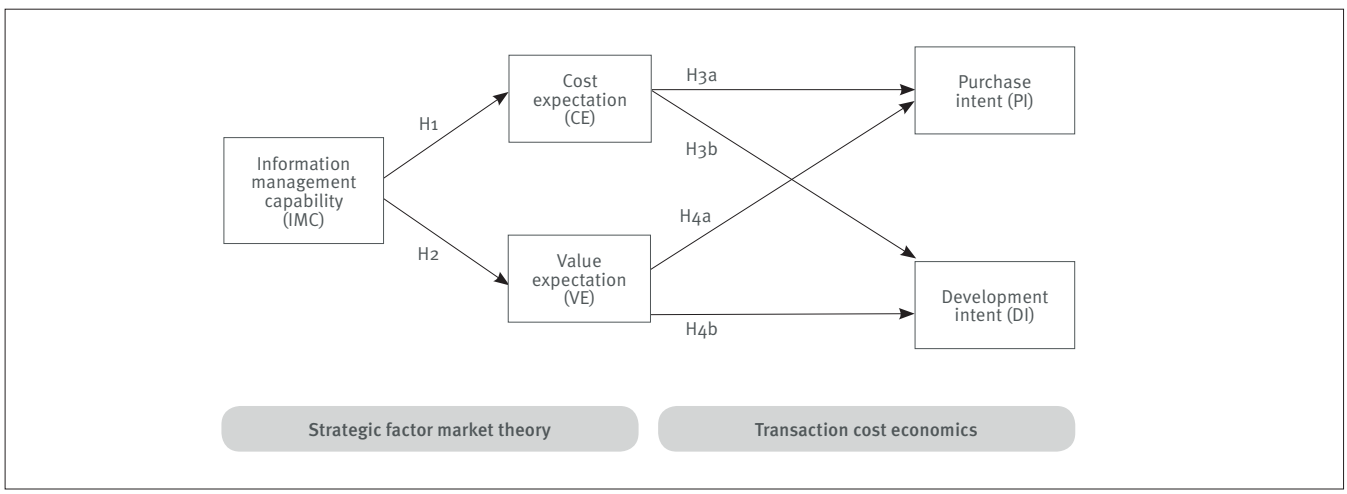
Just as we expect to see companies with better developed IMC to have a lower expectation of the costs necessary to employ a Big Data strategy, it is also likely that this prediction of reduced costs favors a greater predisposition toward implementation. Additionally, with a more accurate cost expectation, companies with an elevated IMC level can create an adequate strategy within their budgets. We also expect the opposite effect: firms with less developed IMC will tend to have less exact cost predictions and therefore greater uncertainty when deciding whether to buy or develop resources to implement a Big Data strategy.

> H4a: Firms with greater expectations of the costs to implement Big Data strategies have less purchase intent for these solutions.

> H4b: Firms with greater expectations of the cost to implement Big Data strategies have less intent to develop these solutions internally.

Considering the four-hypothesis developed above, we built the Research Model. An illustrated presentation of this can be seen in Figure 1.

Figure 1. **Research Model**



RESEARCH METHODOLOGY

We tested the hypotheses utilizing partial least squares structural equation modeling (PLS-SEM) based on survey data. PLS-SEM is frequently recommended for research in management because data in this field often do not adhere to a multi-varied normal distribution, while the models are complex and can still be informative. It is also recommended for smaller samples and models with less prior support (Ringle, Silva, & Bido, 2014; Hair, Hult, Ringle, & Sarstedt, 2013). In light of the involved variables and the nature of this research, we consider the use of this statistical technique appropriate for empirically testing the hypotheses of the conceptual model.

However, we conducted a preliminary stage with a survey and Card Sorting analysis to propose a scale to measure IMC. We describe this stage in the next section, followed by the steps and details about the sample, data collection, and validation.

 CC BY ISSN 0034-7590; eISSN 2178-938X

Card sorting to create an IMC scale

We adapted a scale to measure IMC in the quantitative phase through a survey. This scale was based on existing research instruments (Carmichael et al., 2011; Mithas et al., 2011). The need to construct an IMC scale that could handle this new data environment did not influence the other variables, which already have tested scales.

For the scale, we applied the Optimal Workshop tool to perform a Card Sorting with 10 IT executives. Each online participant took an average of seven minutes to complete. Based on the card sorting results, we reduced the scale from 20 items across five dimensions (people, distribution, access, infrastructure, and information architecture) to 10 items by analyzing a matrix in which we used the cut above 60 percent similarity. To evaluate the dimensions, we used a dendrogram analysis for the best merge method, which often outperforms the actual agreement method when a survey has fewer participants. It makes assumptions about more massive clusters based on individual pair relationships (Optimal WorkShop, 2017). The scores of the cut represent 40 percent of the participants who agree with parts of this grouping. Five dimensions emerged from the group of scale-items assessed by the executives, which were in turn selected from the existing literature. We collected this group through Card Sorting analysis and named them based on the gathered items (people, distribution, access, infrastructure, and information architecture) in line with the authors' analysis of the results from the preliminary stage of the study.

We thus developed the IMC scale for this study. We developed this scale because in-depth research about this construct (Mithas et al., 2011) was validated from an adaptation from pre-existing secondary data, and to incorporate elements addressed in other works (Carmichael et al., 2011). The scales for the other variables of the research tool are adapted from the literature and modified as needed for this study. All items used a seven-point Likert scale (1-Strongly Disagree; 7 – Strongly Agree). We conducted the statistical analysis using the SmartPLS version 3.2.0 software package.

Sample frame and data collection

We collected data through an online research created using the Google Forms platform. Data were collected through social networks, primarily through specific discussion groups about the addressed subjects. Some 29,282 people saw the notices, 208 people clicked on them, and we received 114 completed forms. The answer rate was 59 percent. Among these, we eliminated 13 through three validation questions inserted in the questionnaire to help with data quality control, leaving us with a final sample of 101 forms. Thus, the sample exceeds the minimum of 68 cases, for a power of 0.8 and a medium effect size f2 of 0.15 (Hair et al., 2013) with the variables at a maximum number of two predictors. We calculated this minimum sample using the G*Power 3.1 tool (Faul, Erdfelder, Buchner, & Lang, 2009).

The respondents were managers and executives in IT or other areas related to the implementation of IM strategies. Table 1 summarizes the profiles of the respondent firms, from which we can conclude that the sample is diversified and lightly focused on industry and size, whether through the number of employees or invoicing. The two most apparent differences in the size variable appear in the first two rows. In the first row, there is a smaller percentage of firms invoicing up to one million dollars (16%), while the percentage of companies with up to 50 employees is 27 percent. In contrast, the second row presents a greater percentage of invoicing (23% from 1 to 6.7 million dollars) and a smaller number of employees. A possible explanation for these differences may be in the high number of technological jobs, which have a high profitability potential with fewer employees. There were significant differences in the results relating to industry or firm size. In using Finite Mixture PLS, we did not identify latent classes that evidence the presence of groups within a sample.

Table 1. **Respondent firms' profiles**

Industry	%	Number of employees	%	*Annual revenue*	%
Technology	24%	Up to 50	27%	Up to 1 million dollars	16%
Manufacturing	18%	51 - 100	13%	1 to 6.7 million dollars	23%
Financial services	12%	101 - 500	11%	6.7 to 37.5 million dollars	14%
Professional services	11%	501 - 1,000	16%	37.5 to 125 million dollars	12%
Others	35%	More than 1,000	33%	More than 125 million dollars	36%

Note: n=101

CC BY

ISSN 0034-7590; eISSN 2178-938X

RESULTS

We first present an analysis of the results in terms of the measurement model, followed by an evaluation of the structural model.

Evaluation of the measurement model

We evaluated the measurement model through a series of reliability tests, including composite reliability (CR), Cronbach's alpha, average variance extracted (AVE), and discriminant validity (Hair et al., 2013; Ringle et al., 2014). As Table 2 shows, following Fornell and Larcker's (Henseler, Ringle, & Sinkovics, 2009) criteria, the model converges, and the result is satisfactory because the AVE is above 0.50 for all variables.

Although the traditional indicator to evaluate internal consistency is Cronbach's alpha, CR is the best for PLS-PM because it is the least sensitive to the number of items in each construct (Ringle et al., 2014). In Table 2, we also see that all the variables present both indicators (Cronbach's alpha and CR) above 0.7. Therefore, all the variables are considered adequate and satisfactory (Hair et al., 2013). Also in Table 2, we report the Fornell and Larcker (1981) criteria to verify the discriminant quality according to the correlating values between the variables. The results indicate no correlation between distinct variables greater than the square root of the AVE of each variable (highlighted in gray in the main diagonal).

As the last criterion to evaluate the quality of the measurement model, we calculated discriminant validity utilizing a cross-loading analysis (Chin, 1998). In Table 3 we find no indicators with factor loadings below their variable than in others. Having attended to the quality criteria and discriminant validity of the model, we next evaluate the structural model in the next sub-section.

Table 2. **Quality Criteria**

Variables	AVE	Composite reliability	Cronbach's Alpha	CE	DI	IMC	PI	*VE*
Cost expectation	0.778	0.875	0.715	**0.882**				
Development intent	0.698	0.874	0.784	-0.304	**0.836**			
IMC	0.548	0.923	0.907	-0.407	0.258	**0.740**		
Purchase intent	0.657	0.851	0.747	-0.405	0.735	0.300	**0.811**	
Value expectation	0.819	0.901	0.780	-0.392	0.318	0.647	0.360	**0.905**
Mean				4,75	3,26	4,18	3,40	5,16
SD				1,64	1,87	1,64	1,92	1,67

Note: CE = Cost expectation; DI = Development intent; IMC = Information management capability; PI = Purchase intent; VE = Value expectation.

Table 3. **Cross-Loadings**

Items x Variables	IMC	CE	DI	PI	VE
IMC1	**0.585**	-0.178	0.022	0.004	0.363
IMC2	**0.757**	-0.255	0.236	0.263	0.459
IMC3	**0.784**	-0.273	0.177	0.165	0.543
IMC4	**0.823**	-0.347	0.319	0.351	0.656
IMC5	**0.817**	-0.289	0.190	0.203	0.600
IMC6	**0.697**	-0.182	0.033	-0.048	0.349
IMC7	**0.735**	-0.265	0.308	0.480	0.486
IMC8	**0.686**	-0.293	0.107	0.286	0.425
IMC9	**0.711**	-0.417	0.125	0.191	0.337
IMC10	**0.773**	-0.455	0.259	0.186	0.452
CE1	-0.387	**0.885**	-0.299	-0.316	-0.390
CE2	-0.331	**0.879**	-0.237	-0.399	-0.301
DI1	0.253	-0.285	**0.826**	0.819	0.305
DI2	0.239	-0.204	**0.892**	0.588	0.253
DI3	0.145	-0.261	**0.786**	0.385	0.226
PI1	0.253	-0.285	0.826	**0.819**	0.305
PI2	0.249	-0.404	0.481	**0.858**	0.362
PI3	0.229	-0.269	0.517	**0.751**	0.166
VE1	0.557	-0.361	0.325	0.362	**0.907**
VE2	0.615	-0.349	0.250	0.289	**0.903**

Evaluation of the structural model

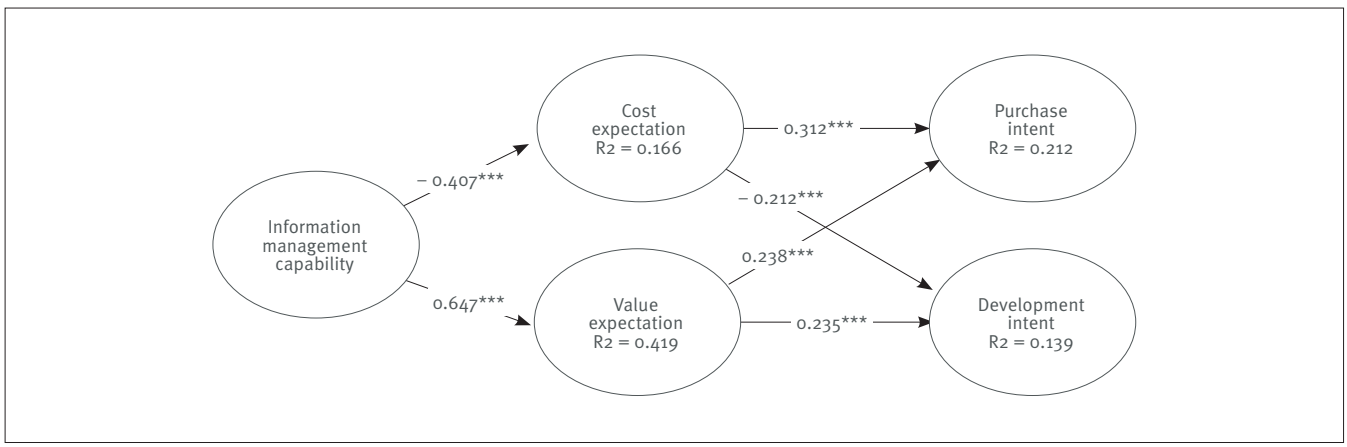
To test the hypotheses and the predictive power of the model, we calculated Pearson's coefficients of determination (R2), the effect size (f2), predictive validity (Q2), and path coefficient (r). According to Cohen's (1988) criteria, we can verify a medium effect of the model on the cost expectation (CE) (0.166) and development intent (DI) (0.139) variables, and a large effect on the value expectation (VE) (0.419) variable, and an almost large effect on the purchase intent (0.212) variable.

The bootstrapping analysis with 1,000 samples demonstrates that all the relations of the observable variables with the latent variables, and those among the latent variables, have significant correlations and regression coefficients at p<0.001, rejecting Ho. We then performed two other quality evaluations of the model adjustment, the predictive validity (Q2) and the effect size (f2), through the blindfolding procedure. Table 4 shows that all Q2s are above zero, demonstrating the model's accuracy. The analysis of the effect size considers a medium utility of CE, DI, and purchase intent (PI) to adjust the model. The results are close to an almost large utility of VE according to the criteria in Hair et al. (2013). Finally, the path coefficients, illustrated in Figure 2, show that the results support all hypotheses.

Table 4. **Results of R^2, Q^2, and f^2**

Relations	R^2	Q^2	f^2
CE	0.166	0.112	0.189
DI	0.139	0.085	0.143
PI	0.212	0.111	0.119
VE	0.419	0.333	0.339

Figure 2. **Results of the empirical model: Path coefficients and R^2**



Note: * p<0.05; ** p<0.01; *** p<0.001

According to the theoretical assumptions of SFM, H1 was confirmed since IMC had a negative impact on the CE of Big Data strategies; that is, the more developed a firm's IMC is, the lower the expectation of the expense to implement a Big Data strategy. The path coefficient analysis highlights that the IMC effect is even more evident on the VE expectation of these strategies. Hypothesis 2 was confirmed, indicating that this ability can be a potential source of imperfections in the SFM for Big Data in both cases.

The other half of the model (H3 and H4) depicts the impact of the expectation to implement Big Data strategies in terms of the cost and value on the intention to purchase (H3 and H4) and to develop (H3b ad H4b) these capabilities. Both hypotheses were confirmed. This impact was negative for Hypotheses 3a (purchase) and 3b (develop), demonstrating that a high cost expectation has a negative impact on the intent to purchase or develop Big Data strategies. The results also confirm Hypotheses 4a and 4b. In other words, the intention to purchase or develop Big Data strategies was positive when the expectation of VE from a Big Data strategy was higher.

FINAL CONSIDERATIONS

We finalize this section with a discussion of the subject and an outline of future research directions.

Contributions to research

This paper contributes to the literature on management information systems by exploring a relatively recent theme (Big Data) and its relation to a firm's existing capability (IMC).

CC BY

ISSN 0034-7590; eISSN 2178-938X

Specifically, we analyzed this phenomenon by focusing on its impact on organizations. "This focus creates a tighter linkage between data and business models: we care deeply about business transformation and value creation through data, and less for algorithms or frameworks without a linkage to business value" (Agarwal & Dhar, 2014, p. 445).

First, the research employed a rare theory in IS – SFM. This theory, along with transaction cost theory (widely used in IS), supported the development of the hypotheses and confirmed the statistical analysis. With this theoretical foundation and from the indications in the literature, it was possible to establish Hypothesis 1. Our results attest that IMC can have a negative impact on the expected cost of the necessary resources to implement a Big Data strategy. These results confirm that organizations have different cost expectations in the search for strategic resources (Barney 1986). IMC plays a relevant role in this heterogeneity of perceptions, whether through more accuracy (Mithas et al., 2011) in the access to and distribution of information, or the perceptive polarization effect (Vasconcelos et al., 2006). Companies that were not successfully able to develop IMC may have a higher expectation of the cost to implement a new strategy related to IM. However, this effect appears to be more strongly evident in the relationships in Hypothesis 2. We demonstrated that IMC positively impacts the expected value extraction from a Big Data strategy. This was the most elevated effect we found, which may indicate a product of the developed abilities or a reflex of successful experiences with IM.

On the other hand, we explained the impact of the expected cost on the intent of purchase or develop the resources and capabilities to implement a strategy to deal with voluminous and heterogeneous data through Hypotheses 3a (purchase) and 3b (develop). The negative impact was supported by empirical data demonstrating that a high cost expectation has an even more negative impact on purchase intent than on the intent to develop the resources and capabilities necessary for the strategy internally. Conversely, the results supported Hypothesis 4 (H4a and H4b), showing that a greater expectation of future value extraction positively impacts the intent to purchase or develop Big Data strategies. In this case, the evidenced size effects for the intent to purchase or develop the required resources for these strategies were very similar. Nevertheless, this study did not aim to evaluate whether or not these expectations correspond to market reality. It is important to note that, in general, investments in IS strategies only reduce transaction costs if the firm consumes fewer resources than the economy generates (Ciborra, 1996).

Through two theoretical perspectives, our research contributes to our understanding of the impact that existing IMC may have on the adoption or non-adoption of new strategies in response to changes in information. More importantly, this study revealed the role of this capability as a potential source of imperfections in the SFM and may be a first step to investigating the role of IMC in the competitive performance of firms.

In addition, along with adopting the perspective of the IMC literature, we propose a new definition that is more in tune with the current context and the IM needs of organizations. We also proposed and validated a new scale to measure this construct.

Implications for practice

We can classify the implications of this study on practice for two types of organizations: those that look for solutions to respond to the environmental changes caused by Big Data and those that offer these solutions. For companies planning to implement Big Data strategies, the results reveal a large variation in the expectations of both the value and cost of the needed resources. This variation may reflect opportunities to search the market for underestimated resources or to incur the risk of acquiring overvalued resources. To reduce these risks and improve performance in the search to exploit these opportunities, our results show that investing in IM not only improves organizational performance (Carmichael et al., 2011; Mithas et al., 2011), it may also help firms evaluate future strategies.

From the other side of market, this work may serve firms that offer the resources and capabilities to implement Big Data strategies some insight into the expectations of their current or potential consumers. Understanding the differences in the perceptions of organizations with different levels of IMC may help firms create an adequate solution and contribute to the success of that solution in IMC development at greater levels for their clients.

Limitations and future research

Our study sample was very heterogeneous, as Table 1 shows, as we collected data non-systematically, and it may, thus, not entirely reflect the population of firms. It is also not possible to identify whether the results apply to a specific group of organizations. We measured the purchase intent and cost expectation constructs using only two indicators, and even though both presented good performance in terms of validity and reliability, it is still one indicator less than recommended.

This research opens the way for new investigations in IS, particularly related to IMC, the context of Big Data, and even new

 CC BY

studies making use of SFM theory. Regarding IMC, we believe that future research may strengthen the strategic role of these capabilities, especially in this Big Data context. Researchers can use SFM to analyze other phenomena in the area and connect it to other theories in the IS literature. The model could hold true for IS strategies in general and can be investigated in the context of other technologies (such as business analytics or business intelligence).

CONCLUSION

This study, despite bringing in quantitative results, is exploratory given the nature of the content analyzed. We aimed to investigate how pre-existing IMC within organizations affects the expectations and intent of these firms in adopting a new IM strategy.

Our results offer insights into the effect on the relations between IMC and cost and future value expectation, in addition to the impact of these expectations on the intent to purchase or develop the needed resources to implement a Big Data strategy. Generally, the results unveiled that IMC positively influences value expectation and negatively influences cost expectation. Value expectation homogeneously and positively impacts the intent to purchase or develop these resources. Finally, cost expectation negatively influences development intent and, even more sharply, the purchase intent of the resources and capabilities for Big Data.

If one key resource for survival in this new environment is the ability to obtain access to more information and to be able to manage this information flow (Cordella, 2006), this research contributes to IS literature by exploring the potential of IMC in this Big Data context. From an academic standpoint, this study tested a less common theory in the literature, which researchers can explore further to analyze IS themes. Lastly, this research can help companies that supply Big Data solutions, as well as firms that intend to invest in strategies to deal with this change in the information environment.

ACKNOWLEDGMENT

The authors are grateful for the financial support provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

REFERENCES

Agarwal, R., & Dhar, V. (2014). Editorial-Big Data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research, 25*(3), 443-448. doi:10.1287/isre.2014.0546

Barney, J. (1986). Strategic factor markets: Expectations, luck, and business strategy. *Management Science, 32*(10), 1231-1241.

Brinkhues, R., Maçada, A., & Casalinho, G. (2014). Information management capabilities: Antecedents and consequences. In *Twentieth Americas Conference on Information Systems*. Savannah, 1-11. Retrieved from https://pdfs.semanticscholar.org/1503/001cb9628f35acd727c4b31b02f613f6523c.pdf

Brown, B., Chui, M., & Manyika, J. (2011). Are you ready for the era of "Big Data"? Retrieved from http://www.t-systems.com/solutions/download-mckinsey-quarterly-/1148544_1/blobBinary/Study-McKinsey-Big-data.pdf [Accessed November 23, 2014].

Carmichael, F., Palacios-Marques, D., & Gil-Pechuan., I. (2011). How to create information management capabilities through web 2.0. *The Service Industries, 31*(10), 1613-1625. doi:10.1080/02642069.2010.485635

Chiang, R. H. L., Grover, V., Liang, T.-P., & Zhang, D. (2018). Special Issue: Strategic value of Big Data and business analytics. *Journal of Management Information Systems, 35*(2), 383-387. doi:10.1080/07421222.2018.1451950

Chin, W. W. (1998). The partial least squares approach for structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295-236). London, UK: Laurence Erlbaum Associates.

Ciborra, C. U. (1996). Teams, markets, and systems: Business innovation and information technology. New York, NY: Cambridge University Press.

Coase, R. H. (1937). The nature of the firm. *Economica, 4*(16), 386-405. doi:10.1111/j.1468-0335.1937.tb00002.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* 2nd ed. New Jersey, NJ: Lawrence Erlbaum Associates.

Cordella, A. (2006). Transaction costs and information systems: Does IT add up? *Journal of Information Technology, 21*(3), 195-202. doi:10.1057/palgrave.jit.2000066

Davenport, T. H. (2014). *Big Data at work: Dispelling the myths, uncovering the opportunities*. Boston, MA: Harvard Business School Publishing.

Davenport, T. H., Barth, P., & Bean, R. (2012). How "Big Data" is different. *Sloan Management Review, 54*(1), 21–24.

Davenport, T. H., & Bean, R. (2018). Data and innovation: How Big Data and AI are driving business innovation. New Vantage Partners LLC. Retrieved from https://newvantage.com/wp-content/uploads/2018/01/Big-Data-Executive-Survey-2018-Findings.pdf

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. doi:10.3758/BRM.41.4.1149

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50. doi:10.2307/3151312

Gartner. (2014). Newsroom Gartner survey reveals that 73 percent of organizations have invested or plan to invest in Big Data in the next two years. Retrieved from http://www.gartner.com/newsroom/id/2848718

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information Systems, 26*(3), 191-209. doi:10.1016/j.jsis.2017.07.003

Hair Jr, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage Publications.

Henseler, J., Ringle, C., & Sinkovics, R. (2009). The use of partial least squares path modeling in international marketing. Retrieved from https://opus.lib.uts.edu.au/research/handle/10453/10057

International Data Corporation. (2018). Revenues for Big Data and business analytics solutions forecast to reach $260 billion in 2022, Led by the banking and manufacturing industries, according to IDC. Retrieved from https://www.idc.com/getdoc.jsp?containerId=prUS44215218

Johnson, J. (2012). Big Data + Big Analytics = Big Opportunity. *Financial Executive, 28*(6), 50–53.

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management, 34*(3), 387-394. doi:10.1016/j.ijinfomgt.2014.02.002

Lakoju, M., & Serrano, A. (2017). Framework for aligning Big-Data strategy with organizational goals. Proceedings of Twenty-third Americas Conference on Information Systems, Boston.

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The management revolution. *MIT Sloan Management Review, 90*(10), 61–68.

Mckendrick, J. (2013). 2013 Big Data opportunities survey. *Unisphere Research*. New Providence. http://www.unisphereresearch.com/Issues/5375-Big-Data-Big-Challenges-Big-Opportunities-2012-IOUG-Big-Data-Strategies-Survey.htm

Mithas, S., Lee, M., Earley, S., Murugesan, S., & Djavanshir, R. (2013). Leveraging Big Data and business analytics. *IT Professional, 15*(6), 18-20. doi:10.1109/MITP.2013.95

Mithas, S., Ramasubbu, N., & Sambamurthy, V. (2011). How information management capability influences firm performance. *MIS Quarterly, 35*(1), 237-256.

Mohanty, S., Jagadeesh, M., & Srivatsa, H. (2013). The new information management paradigm. In S. Mohanty, M. Jagadeesh, & H. Srivatsa, *Big Data Imperatives* (pp. 25-44). Berkeley, CA: Apress.

Optimal WorkShop (2017). Card sorting 101. Your guide to creating and running an effective card sort. Retrieved from https://www.optimalworkshop.com/101/card-sorting

Phadtare, M. (2011). *Strategic management: Concepts and cases*. New Delhi, India: PHI Learning Pvt. Ltd.

Ringle, C. M., Silva, D., & Bido, D. de S. (2014). Modelagem de equações estruturais com utilização do Smartpls. *Revista Brasileira de Marketing, 13*(2), 56-73. doi:10.5585/remark.v13i2.2717

Tabesh, P., Mousavidin, E., & Hasani, S. (2019). Implementing big data strategies: A managerial perspective. *Business Horizons, 62*(3), 347–358. doi:10.1016/j.bushor.2019.02.001

Vasconcelos, I. F. G. De, Mascarenhas, A. O., & Vasconcelos, F. C. De, (2006). Gestão do paradoxo "passado versus futuro": Uma visão transformacional da gestão de pessoas. *RAE-Eletrônica, 5*(1). Retrieved from https://rae.fgv.br/rae-eletronica/

Wang, Y., Kung, L. A., Wang, W. Y. C., & Cegielski, C. G. (2018). An integrated big data analytics-enabled transformation model: Application to health care. *Information and Management, 55*(1), 64-79. doi:10.1016/j.im.2017.04.001

White, M. (2012). Digital workplaces: Vision and reality. *Business Information Review, 29*(4), 205-214. doi:10.1177/0266382112470412

FORUM

Submitted 09.26.2018. Approved 07.19.2019
Evaluated through a double-blind review process. Guest Scientific Editors: Eduardo de Rezende Francisco, José Luiz Kugler, Soong Moon Kang, Ricardo Silva, and Peter Alexander Whigham
Original version
DOI: http://dx.doi.org/10.1590/S0034-759020190605

INTENTION TO ADOPT BIG DATA IN SUPPLY CHAIN MANAGEMENT: A BRAZILIAN PERSPECTIVE

Intenção de adoção de big data na cadeia de suprimentos: Uma perspectiva brasileira

Intención de adopción de big data en la cadena de suministros: Una perspectiva brasileña

ABSTRACT

Big data applications have been remodeling several business models and provoking strong radical transformations in supply chain management (SCM). Supported by the literature on big data, supply chain management, and the unified theory of acceptance and use of technology (UTAUT), this study aims to evaluate the variables that influence the intention of Brazilian SCM professionals to adopt big data. To this end, we adapted and validated a previously developed UTAUT model. A survey of 152 supply chain respondents revealed that facilitating conditions (e.g., IT infrastructure) have a high influence on their intention to adopt big data. However, social influence and performance expectancy showed no significant effect. This study contributes to the practical field, offering valuable insights for decision-makers considering big data projects. It also contributes to the literature by helping minimize the research gap in big data in the Brazilian context.

KEYWORDS | Big data, supply chain management, adoption, survey, partial least squares structural equation modeling.

RESUMO

As aplicações de big data têm remodelado vários modelos de negócios e provocado grandes transformações na gestão da cadeia de suprimentos (GCS). Apoiado pela literatura emergente de big data, GCS e teoria unificada de aceitação e uso de tecnologia (UTAUT), este estudo tem como objetivo avaliar as variáveis que influenciam os profissionais brasileiros que atuam na GCS a adotar big data. Assim, nós adaptamos e validamos um modelo UTAUT previamente desenvolvido. Um total de 152 profissionais que atuam na gestão de cadeias de suprimentos revelou que condições facilitadoras (como a infraestrutura de TI) têm uma grande influência na adoção de big data. Por outro lado, a influência social e a expectativa de desempenho não apresentaram efeito significativo. Este estudo contribui para a prática, com conhecimentos valiosos para os tomadores de decisão que estão considerando projetos de big data. Além disso, ele ajuda a minimizar a lacuna em relação aos estudos de big data no contexto brasileiro.

PALAVRAS-CHAVE | *Big data, gestão da cadeia de suprimentos, adoção, survey, partial least squares structural equation modeling.*

RESUMEN

Las aplicaciones de big data han estado remodelando varios modelos de negocios y han provocado fuertes transformaciones en la cadena de suministro (CS). Con el apoyo de la literatura de big data, CS y la teoría unificada de aceptación y uso de la tecnología (UTAUT), este estudio tiene objetivo evaluar las variables que afectan a los profesionales brasileños para adoptar big data. Por lo tanto, adaptamos y validamos un modelo UTAUT previamente desarrollado. Un total de 152 encuestados de CS revelaron que las condiciones de facilitación (por ejemplo, la infraestructura de TI) tienen una gran influencia en la adopción de big data. Por otro lado, la influencia social y la expectativa de desempeño no mostraron un efecto significativo. Este estudio contribuye a la práctica, con información valiosa para los responsables de la toma de decisiones que están considerando proyectos de big data. Además, ayudamos a minimizar la brecha con respecto a los estudios de big data en el contexto brasileño.

PALABRAS CLAVE | *Big data, gestión de la cadena de suministro, adopción, survey, partial least squares structural equation modeling.*

MACIEL M. QUEIROZ[1]
maciel.queiroz@docente.unip.br
ORCID: 0000-0002-6025-9191

SUSANA CARLA FARIAS PEREIRA[2]
susana.pereira@fgv.br
ORCID: 0000-0002-3952-7489

[1]Universidade Paulista, Programa de Pós-graduação em Administração, São Paulo, SP, Brazil

[2]Fundação Getulio Vargas, Escola de Administração de Empresas de São Paulo, SP, Brazil

 CC BY

INTRODUCTION

The rapid advancement of information and communication technologies (ICTs) has motivated logistics and supply chain management practitioners and scholars (Zinn & Goldsby, 2017b, 2017a) to understand the role of these new technologies, and to determine how organizations can capture value through ICT adoption. A highly disruptive and significant technology that has emerged recently is big data (Davenport, 2006; Manyika et al., 2011; Rotella, 2012). The amount of data produced everyday has been increasing drastically (Domo, 2017). This growth has imposed several complexities concerning its management. In this context, big data offers a powerful approach to helping organizations analyze (Croll, 2015) large amounts of data to provide insights into the decision-making process (Abawajy, 2015).

The literature considered big data the "next big thing in innovation" (Gobble, 2013, p. 64) and "the fourth paradigm of science" (Strawn, 2012, p. 34). Big data has impacted practically all business models. For instance, 35% of Amazon.com's revenue is generated through the use of big data (Wills, 2014), along with the remodeling of marketing activities that capture rich data on consumer behavior in real-time (Erevelles, Fukawa, & Swayne, 2016). A field that has been making substantial efforts to harness big data is supply chain management (SCM) (Gunasekaran et al., 2017; Kache & Seuring, 2017; Richey, Morgan, Lindsey-Hall, & Adams, 2016; K. J. Wu et al., 2017; R. Zhao, Liu, Zhang, & Huang, 2017).

Despite the potential benefits of employing big data in supply chain management (Hazen, Boone, Ezell, & Jones-Farmer, 2014; Kache & Seuring, 2017; Schoenherr & Speier-Pero, 2015), awareness of and initiatives on big data in the Brazilian SCM context are rare, and the literature lacks strong empirical results (Queiroz & Telles, 2018). The current initial stage of big data adoption presents an opportunity for scholars and practitioners to fill this gap. For example, to the best of our knowledge, no previous study analyzed the intention of Brazilian SCM professionals to adopt big data. To bridge this gap, this study provides an in-depth understanding of Brazilian supply chain professionals' intention to use big data. We adapt a previously developed and validated unified theory of acceptance and use of technology (UTAUT) model (Venkatesh, Morris, Davis, & Davis, 2003; Queiroz & Wamba, 2019), by including a trust construct. More specifically, this study answers the following research question: How do the variables from the UTAUT model explain Brazilian SCM professionals' intention to adopt big data?

To answer this question, this work draws on the literature on big data (Davenport, 2006; Manyika et al., 2011; Queiroz & Telles, 2018), supply chain management (Carter, Rogers, & Choi, 2015; Mentzer et al., 2001), and UTAUT (Venkatesh et al., 2003; Venkatesh, Thong, & Xu, 2012; Queiroz & Wamba, 2019) to develop the hypotheses and model. The conceptual model was adapted and validated with partial least squares structural equation modeling (PLS-SEM). The main findings offer strong theoretical and managerial implications. From the managerial perspective, we verified that facilitating conditions (e.g., infrastructure) exert high influence on the behavioral intention of big data adoption. From the theoretical lens, our findings revealed that neither social influence nor performance expectancy are good predictors of the behavioral intention of big data adoption in Brazilian SCM professionals.

This paper is organized as follows: next, we present the emerging theoretical foundations for big data research, SCM, and UTAUT. Then, the hypotheses and the research model are described, followed by the survey methodology and analysis using PLS-SEM. That is succeeded by a discussion on managerial and theoretical implications as well as limitations of the current work and directions for future research. Finally, our conclusions are elucidated.

THEORETICAL BACKGROUND

Big Data: Fundamentals, concepts, and challenges

Big Data has emerged as a highly disruptive information and communication technology (ICT). A well-articulated and suitable definition of Big Data is "[...] datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze" (Manyika et al., 2011, p. 1). Thus, Big Data can be regarded as providing a robust approach to exploring data in the context of descriptive, prescriptive, and predictive decisions (Phillips-Wren & Hoskisson, 2015). This approach is commonly called Big Data analytics (BDA), and is represented by a 5V approach (volume, velocity, variety, veracity, and value) (Queiroz & Telles, 2018; Wamba et al., 2017). In other words, BDA uses sophisticated statistics, mathematical and computational techniques to explore a large set of data to provide insights to decision-makers. In this study, we use the definition of Big Data proposed by Phillips-Wren and Hoskisson (2015).The authors described Big Data as data that overtake the organization's capabilities, regarding storage, and analysis to support and bring insights to the decision-making process.

The volume of data has increased drastically in recent years, mainly because of the variety of data produced today (Bibri &

 CC BY ISSN 0034-7590; eISSN 2178-938X

Krogstie, 2017) (e.g., ERP systems, Twitter, Facebook, Google, Linkedin, GPS, among others) and the velocity of its spread (Munshi & Mohamed, 2017; Srinivasan & Swink, 2018). This complex scenario impels organizations to develop distinctive capabilities for storing, processing, and analyzing data to support the decision-making process. However, creating value is not a trivial task, mainly because of organizations' limited capacity to process and analyze a variety of data. Moreover, data veracity, which indicates data quality and trustworthiness (Munshi & Mohamed, 2017; Nobre & Tavares, 2017), seems to be a huge challenge for organizations.

In the SCM-related fields, Big Data is being newly explored in different contexts: in SCM agility enhancement with Big Data and multi-agent-based systems (Giannakis & Louis, 2016), in an optimization of green SCM considering hazardous materials and carbon emission (R. Zhao et al., 2017), in the manufacturing sector (Zhong, Newman, Huang, & Lan, 2016), and in the information exploitation of SCM (Kache & Seuring, 2017). It is clear that Big Data can improve organizations' performance significantly (Akter, Wamba, Gunasekaran, Dubey, & Childe, 2016; Gunasekaran et al., 2017; Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015; G. Wang, Gunasekaran, Ngai, & Papadopoulos, 2016).

Supply chain management and the impacts of cutting-edge technologies

Recently, the logistics and SCM fields have been significantly impacted by the exponential growth in ICT usage. Accordingly, scholars and practitioners have strived to understand its potential effects and application opportunities in their business models (Zinn & Goldsby, 2017a, 2017b). In this context, SCM is defined as:

> The management of a network of relationships within a firm and between interdependent organizations and business units consisting of material suppliers, purchasing, production facilities, logistics, marketing, and related systems that facilitate the forward and reverse flow of materials, services, finances and information from the original producer to final customer with the benefits of adding value, maximizing profitability through efficiencies, and achieving customer satisfaction (Stock & Boyer, 2009, p. 706).

Moreover, SCM can be viewed as a network (Carter et al., 2015) as well as a complex adaptive system (Choi, Dooley, & Rungtusanatham, 2001), and this complexity has impacted the increasing amount of data. Considering the use of Big Data in SCM, it is clear that it assists in the decision-making process by providing powerful insights into SCM dynamics (e.g., customer buying patterns, cost analysis, market trends). With the help of robust prescriptive and descriptive analysis (G. Wang et al., 2016), businesses have witnessed many cases of significant performance enhancement (Akter et al., 2016; Gunasekaran et al., 2017).

Technology acceptance models (TAMs) and Unified theory of acceptance and use of technology (UTAUT)

Scholars have studied the diffusion and proliferation of information technology (IT) (Davis, 1989; Wamba, 2018; Morris & Venkatesh, 2000; Venkatesh & Brown, 2001) in terms of individuals' beliefs and behavior toward their adoption and use (Mamonov & Benbunan-Fich, 2017; Youngberg, Olsen, & Hauser, 2009). The technology acceptance model (TAM) is a seminal and influential contribution in technology adoption (Davis, 1989), with its roots in the theory of reasoned action (TRA) (Azjen & Fishbein, 1980). The core of the TAM resides in two latent variables: perceived usefulness (PU) and perceived ease of use (PEOU). More recently, Venkatesh et al. (2003) proposed the consolidation of the acceptance model theories leading previously into the unified theory of acceptance and use of technology (UTAUT).

UTAUT

The UTAUT model (Venkatesh et al., 2003) is a robust and influential approach to understanding technology adoption and use at the individual behavior level. The model has four constructs directly focused on technology's intended use: performance expectancy, effort expectancy, social influence, and facilitating conditions.

Performance expectancy refers to "the degree to which an individual believes that using the system will help him or her to attain gains in job performance" (Venkatesh et al., 2003, p. 447). Effort expectancy is "the degree of ease associated with the use of the system" (Venkatesh et al., 2003, p. 450). Social influence denotes "the degree to which an individual perceives that important others believe he or she should use the new system" (Venkatesh et al., 2003, p. 451). Finally, facilitating conditions indicates "the degree to which an individual believes that an organizational and technical infrastructure exists to

 CC BY ISSN 0034-7590; eISSN 2178-938X

support use of the system" (Venkatesh et al., 2003, p. 453). The UTAUT model also has four moderators: gender, age, experience, and voluntariness of use. However, following a previous study (Weerakkody, El-Haddadeh, Al-Sobhi, Shareef, & Dwivedi, 2013), we do not use these moderators in our adapted model (explained in the next section) because this is a preliminary study of BDA adoption in the Brazilian SCM context.

Hypotheses and research model

Supported by the emerging literature on Big Data, SCM, and UTAUT, we adapted a recent model reported in Queiroz and Wamba (2019) to comprehend the Big Data adoption behavior of Brazilian supply chain professionals. We adopted some of the constructs and hypotheses proposed in Queiroz and Wamba´s (2019) model (Figure 1) as these have been adopted and validated by previous studies (Exhibit 1). To these previous constructs reported in Queiroz & Wamba (2019) we added a trust construct, previously validated in the literature (Alalwan, Dwivedi, & Rana, 2017; Gefen, Karahanna, & Straub, 2003). Moreover, the constructs in our model have different relationships than the ones reported in the literature (Queiroz & Wamba, 2019).

Facilitating conditions

Facilitating conditions play a fundamental role in predicting user acceptance and usage behavior (Venkatesh et al., 2003, 2012). In this study, facilitating conditions denotes SCM professionals' knowledge of their organization's capabilities and infrastructure available to support the use of Big Data. Previous studies have reported that facilitating conditions are a good predictor of the behavioral intention of Big Data adoption (Huang, Liu, & Chang, 2012; Sabi, Uzoka, Langmia, & Njeh, 2016). In this study, we theorize that facilitating conditions, besides influencing behavioral intention directly, are critical in professionals' effort expectancy (Dwivedi et al., 2017) and influence their performance expectancy (C. Wang, Jeng, & Huang, 2017). Therefore, we propose the following hypotheses:

> H1a: Facilitating conditions positively affects effort expectancy.
>
> H1b: Facilitating conditions positively affects performance expectancy.
>
> H1c: Facilitating conditions positively affects behavioral intention to adopt Big Data.

Trust

The trust construct has been studied extensively in the business management and management information systems (MIS) fields (Colquitt & Rodell, 2011; K. Wu, Zhao, Zhu, Tan, & Zheng, 2011). Trust is defined as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (Mayer, Davis, & Schoorman, 1995, p. 712). This definition implies that trust is a willingness to depend on the partner based on integrity, benevolence, and credibility. In this context, Big Data is trustworthy for users. In line with prior works (K. Wu et al., 2011), we hypothesize that:

> H2a: Trust positively affects performance expectancy.
>
> H2b: Trust positively affects behavioral intention to adopt Big Data.

Social influence

As reported previously, social influence is a good predictor of technology behavioral intention and usage (Venkatesh et al., 2003). In this work, social influence denotes the extent to which SCM professionals believe their colleagues should use Big Data. Previous studies highlight social influence as a predictor of behavioral intention (Batara, Nurmandi, Warsito, & Pribadi, 2017; Oliveira, Faria, Thomas, & Popovič, 2014; Venkatesh et al., 2012). Our study argues that in the SCM context, social influence relationships exert significant influence on trust (A. Chin, Wafa, & Ooi, 2009) and, in turn, on the behavioral intention (Alalwan et al., 2017). Thus, we propose the following hypotheses:

> H3a: Social influence positively affects trust.
>
> H3b: Social influence positively affects behavioral intention to adopt Big Data.

Effort expectancy

Effort expectancy is related to the system's complexity of operation (Venkatesh et al., 2003). In this study, effort expectancy refers to the ease of use of Big Data systems for an SCM professional. Previous studies discussed the direct effect of effort expectancy in the behavioral intention and usage of a new technology (Batara et al., 2017; Venkatesh et al., 2012; Y. Zhao, Ni, & Zhou, 2018) as well as in the adoption of blockchain in the SCM field (Francisco & Swanson, 2018). Accordingly, this study hypothesizes that:

 CC BY ISSN 0034-7590; eISSN 2178-938X

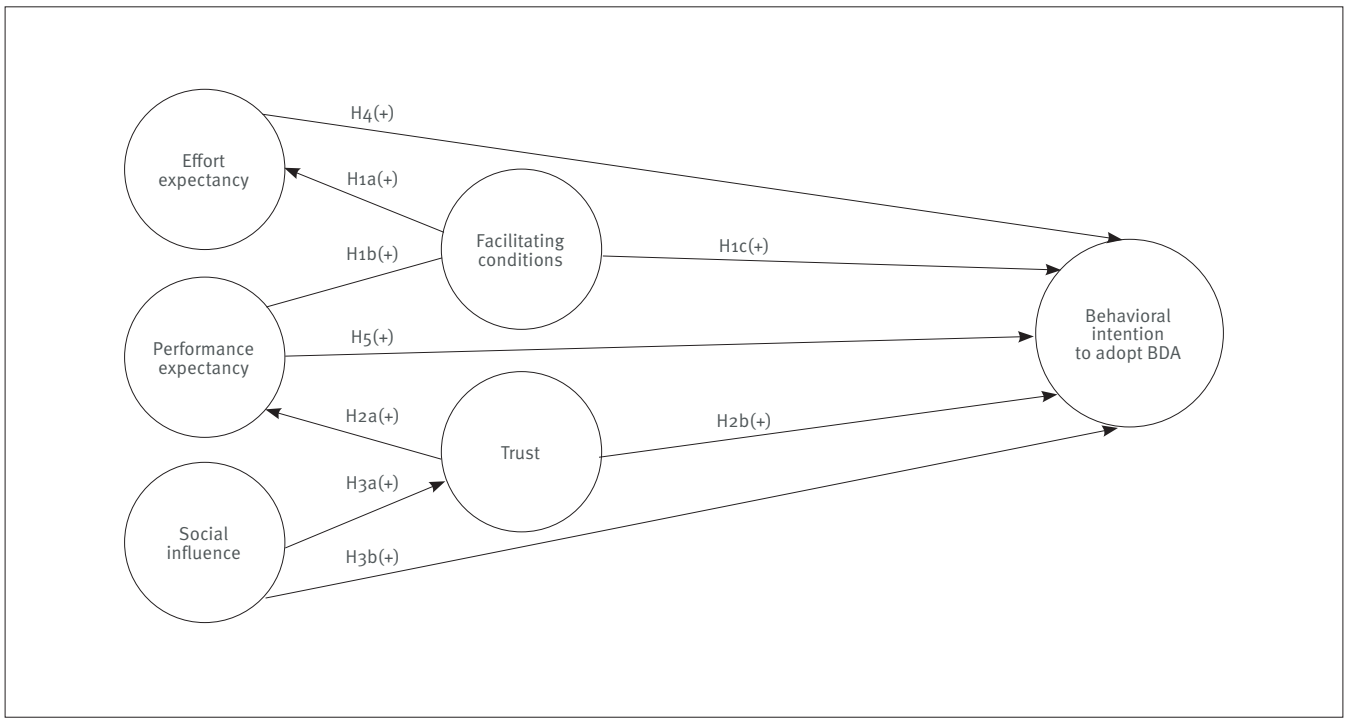
H4: Effort expectancy positively affects behavioral intention to adopt Big Data.

Performance expectancy

In this work, performance expectancy denotes the level to which an SCM professional perceives that Big Data will improve his productivity and performance. With Big Data application, organizations can analyze different types of data employing powerful statistics and machine learning techniques (Kune, Konugurthi, Agarwal, Chillarige, & Buyya, 2016). This implies considerable time savings and productivity improvement for organizations, therefore helping enhance its performance (Gunasekaran et al., 2017; Wamba et al., 2017). Thus, we propose that:

H5: Performance expectancy positively affects behavioral intention to adopt Big Data.

Figure 1. **Conceptual model**



METHODOLOGY

Sample and data collection

A survey instrument based on Queiroz and Wamba (2019) was used to test our proposed hypotheses. The web-based questionnaire was grounded on constructs and scales that have been validated by previous studies (Venkatesh et al., 2003, 2012; Gefen et al., 2003). The Queiroz and Wamba (2019) model was developed based on previous studies; their constructs were adapted from recent studies on TAMs (Alalwan et al., 2017; Venkatesh et al., 2003, 2012). As our main objective was to identify the intention to adopt Big Data, we adapted the Queiroz and Wamba (2019) survey instrument. All constructs were measured using a seven-point Likert scale [1 (strongly disagree) to 7 (strongly agree)] (Wamba et al., 2017). Before data collection, a pilot test was performed with five senior academics and five senior SCM professionals. Data were collected through the LinkedIn social network (Gupta & George, 2016; Queiroz & Telles, 2018). After the pilot, we sent the questionnaire to 600 Brazilian supply chain professionals with experience in Big Data. The survey was conducted in August 2018, and a total of 152 questionnaires were received, representing a response rate of 25.33%. Exhibit 1 shows the constructs and their respective items. We validated the questionnaire by employing outer loadings (Hair et al., 2017), Cronbach's alpha, composite reliability, average variance extracted (Hair et al., 2017; Nunnally, 1978; Riffai, Grant, & Edgar, 2012), and discriminant validity.

 CC BY ISSN 0034-7590; eISSN 2178-938X

Exhibit 1. **Measurement items**

Construct	Label	Items	Sources
Performance expectancy (PE)	PE1	I find big data useful in my daily life.	(Alalwan et al., 2017; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019)
	PE2	Using big data increases my chances of completing tasks that are important to me.	
	PE3	Using big data helps me accomplish tasks more quickly.	
	PE4	Using big data increases my productivity.	
Effort expectancy (EE)	EE1	Learning how to use big data is easy for me.	(Alalwan et al., 2017; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019)
	EE2	My interaction with big data is clear and understandable.	
	EE3	I find big data easy to use.	
	EE4	It is easy for me to become skilful at using big data.	
Social influence (SI)	SI1	People who are important to me think I should use big data.	(Alalwan et al., 2017; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019)
	SI2	People who influence my behaviour think I should use big data.	
	SI3	People whose opinions that I value prefer I use big data.	
Facilitating conditions (FC)	FC1	I have the resources necessary to use big data.	(Alalwan et al., 2017; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019)
	FC2	I have the knowledge necessary to use big data.	
	FC3	Big data is compatible with other technologies I use.	
	FC4	I can get help from others when I have difficulties using big data.	
Behavioral intention to use (BI)	BI1	I intend to use big data in the future.	(Alalwan et al., 2017; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019)
	BI2	I expect to use big data in the future	
	BI3	I plan to use big data in future.	
Trust (TR)	TR1	I believe that big data is trustworthy.	(Alalwan et al., 2017; Gefen et al., 2003)
	TR2	I have trust in big data.	
	TR3	I do not doubt the honesty of big data.	
	TR4	I feel assured that legal and technological structures adequately protect me from problems on big data.	
	TR5	Big data has the ability to fulfil its task.	

 CC BY ISSN 0034-7590; eISSN 2178-938X

RESULTS AND ANALYSIS

Partial least squares structural equation modeling (PLS-SEM) (Ringle, Wende, & Becker, 2015; Shim, Lee, & Kim, 2018; Sun & Teng, 2017) was applied to analyze the research model. PLS-SEM is a powerful approach for analyzing simple and robust models in business management (Hair, Sarstedt, Hopkins, & Kuppelwieser, 2014; Hair, Hult, Ringle, & Sarstedt, 2017), and has gained the attention of SCM scholars (Autry, Williams, & Golicic, 2014; Grawe, Daugherty, & Ralston, 2015; Han, Wang, & Naim, 2017; Yadlapalli, Rahman, & Gunasekaran, 2018). Its main advantages are its flexibility in working with small samples (e.g., 100 respondents) and its formative and reflective constructs (Hair et al., 2017).

Table 1 reports the characteristics of the respondents. Male respondents comprised almost 90% of the total. Regarding age distribution, most respondents (52.63%) were aged 34-41 years. A total of 55.26% respondents had a postgraduate/MBA—the highest education level in our sample—followed by 39.47% holding bachelor degrees and 5.26% holding a master of science degree. Considering the experience at their respective organizations, 50% respondents had worked there for 2-5 years, followed by 21.05% having worked for 6-10 years and 18.42% working for less than one year. Finally, 46.05% of the sample comprised logistics analysts, followed by 26.32% transportation managers, 18.42% operations managers, and 9.21% supply chain managers.

We analyzed the research model with SmartPLS 3.0 (Hair et al., 2017; Ringle et al., 2015). First, the model was assessed by its loadings, Cronbach's alpha, composite reliability, average variance extracted, and discriminant validity.

Measurement model

All outer loadings highlighted in Table 2 exceeded the 0.70 threshold recommended in the literature (Hair et al., 2017). Table 3 shows the main measures for construct reliability and internal consistency of items. In this study, both Cronbach's alpha value and composite reliability were above the 0.70 threshold, and all average variance extracted values were above the 0.50 threshold (Hair et al., 2017; Nunnally, 1978; Riffai, Grant, & Edgar, 2012). Therefore, all constructs in the model have their utilization validated. Table 4 presents the discriminant validity results. In this case, the square root of the average variance extracted for each construct should be greater than the correlations between the constructs (Fornell & Larcker, 1981; Henseler, Ringle, & Sinkovics, 2009). Our results are higher than the 0.70 threshold (Fornell & Larcker, 1981), confirming that all constructs show discrimination (Ahmad & Khalid, 2017; Martins, Oliveira, & Popovič, 2014).

Table 1. **Demographic profile of respondents (n=152)**

Gender	**n**	**%**
Male	136	89.5
Female	16	10.5
Age		
26-33	40	26.32
34-41	80	52.63
42-49	12	7.89
50+	20	13.16
Highest education level		
Bachelor degree	60	39.47
Postgraduate/MBA	84	55.26
Master of Science (MSc)	8	5.26
Number of years spent working in the organization		
Less than one year	28	18.42
2-5 years	76	50.00
6-10 years	32	21.05
11-15 years	16	10.53
Occupation		
Logistics analyst	70	46.05
Operations manager	28	18.42
Transportation manager	40	26.32
Supply chain manager	14	9.21

 CC BY © RAE | São Paulo | **59(6)** | **November-December 2019** | 389-401 ISSN 0034-7590; eISSN 2178-938X

Table 2. **Factor loadings**

	BI	EE	FC	PE	SI	TR
BI1	0.887					
BI2	0.916					
BI3	0.900					
EE1		0.945				
EE2		0.906				
EE3		0.947				
EE4		0.900				
FC1			0.789			
FC2			0.884			
FC3			0.719			
FC4			0.800			
PE1				0.764		
PE2				0.803		
PE3				0.914		
PE4				0.899		
SI1					0.953	
SI2					0.983	
SI3					0.967	
TR1						0.964
TR2						0.944
TR3						0.918
TR4						0.937
TR5						0.913

Note: BI = Behavioral intention
EE = Effort expectancy
FC = Facilitating conditions
PE = Performance expectancy
SI = Social influence
TR = Trust.

Table 3. **Reliability measures**

Construct	Cronbach's alpha	Composite reliability	Average variance extracted
BI	0.881	0.926	0.808
EE	0.942	0.959	0.853
FC	0.806	0.873	0.637
PE	0.864	0.908	0.715
SI	0.965	0.977	0.934
TR	0.963	0.971	0.871

 CC BY ISSN 0034-7590; eISSN 2178-938X

Table 4. **Discriminant validity**

Construct	BI	EE	FC	PE	SI	TR
BI	**0.901**					
EE	0.527	**0.925**				
FC	0.614	0.577	**0.800**			
PE	0.314	0.238	0.620	**0.847**		
SI	0.399	0.390	0.508	0.421	**0.968**	
TR	0.511	0.456	0.580	0.640	0.715	**0.935**

Structural model

Table 5 and 6 present the results of our structural model. Table 5 highlights the path coefficients statistics.The findings indicated that FC has a significant positive effect on EE (β = 0.578, p ‹ 0.001). Thus, H1a is supported. H1b hypothesized that FC has a significant positive effect on PE. The results (β = 0.380, p ‹ 0.001) support H1b. H1c theorized that FC has a significant positive effect on BI. This hypothesis was also supported (β = 0.490, p ‹ 0.001). Next, H2a argued that TR has a significant positive effect on PE. Our results (β = 0.413, p ‹ 0.001) support this hypothesis. Then, H2b argued that TR has a significant positive effect on BI. The results supported H2b (β = 0.327, p ‹ 0.05). H3a theorized that SI has a significant positive effect on TR. The results supported H3a (β = 0.710, p ‹ 0.001). The rest of the hypotheses had unexpected results. H3b theorized that SI has a significant positive effect on BI. Surprisingly, the relationship was found to be negative and non-significant. Therefore, H3b was not supported (β = -0.073, p = 0.519). H4 argued that EE has a significant positive effect on BI. This hypothesis was not supported either (β = 0.166, p ‹ 0.1). Next, H5 theorized that PE has a significant positive effect on BI. Surprisingly, the results (β = -0.214, p ‹ 0.05) showed a negative significant effect on BI. Thus, H5 was not supported.Table 6 demonstrates the variance of the model: 46% variance in BI; 33.30% in EE; 49.80% in PE; and finally, 50.30% in TR. In line with the literature (W. W. Chin, 1998), all r-squares of the model exceeded the 0.20 threshold (Martins et al., 2014).

Table 5. **Path coefficients**

Path	Beta	Standard deviation	t-statistics	p-value	Result
FC -› EE	0.578	0.053	10.921*	0.000	Supported
FC -› PE	0.380	0.064	5.875*	0.000	Supported
FC -› BI	0.490	0.097	5.016*	0.000	Supported
TR -› PE	0.413	0.080	5.301*	0.000	Supported
TR -› BI	0.327	0.112	2.987**	0.003	Supported
SI -› TR	0.710	0.047	15.17*	0.000	Supported
SI -› BI	-0.073	0.097	0.646	0.519	Rejected
EE -› BI	0.166	0.090	1.86	0.063	Rejected
PE -› BI	-0.214	0.099	2.184**	0.029	Rejected

Note: *p ‹ 0.001; **p ‹ 0.05.

 CC BY ISSN 0034-7590; eISSN 2178-938X

Table 6. **R^2 results (dependent variables)**

Construct	R^2	R^2 adjusted
BI	0.477	0.460
EE	0.337	0.333
PE	0.504	0.498
R	0.506	0.503

DISCUSSION AND IMPLICATIONS

The main objective of this study was to gain an in-depth understanding of the intention of Big Data adoption in the Brazilian supply chain context. In light of the lack of Brazil-based studies on cutting-edge technologies (Queiroz and Telles, 2018), this work contributes to advancing the literature on BDA, SCM, and TAMs. The results offer significant managerial and theoretical implications as well as valuable directions to adapt and extend the adopted model.

Managerial implications

We believe the main findings of this study provide important implications for managers and practitioners interested in gaining deeper insights about BDA in SCM and their adoption enablers. In line with the literature that regards Big Data as an essential tool to improve supply chain performance (Gunasekaran et al., 2017; Hazen, Skipper, Ezell, & Boone, 2016; G. Wang et al., 2016), our study first showed that Big Data can be a suitable tool to help SCM managers gain insights and thus support their decision-making process. Second, facilitating conditions exert a high influence on Big Data adoption. This implies that managers have to pay sufficient attention to IT infrastructure, internet speed, and integration with other systems, among other considerations (Sabi et al., 2016; Venkatesh et al., 2003).

Surprisingly, despite the literature reporting performance expectancy as a good predictor of behavioral intention towards technology adoption (Dwivedi et al., 2017; Farooq et al., 2017; Venkatesh et al., 2003; Weerakkody et al., 2013), in this study, performance expectancy was not found to be a good predictor of the behavioral intention to use Big Data among Brazilian SCM professionals. This finding indicates a challenge for managers because it can be a significant barrier to the adoption of Big Data technologies. It also opens up research directions for scholars and practitioners to investigate. On the other hand, social influence as a predictor of trust is a high influencer (A. Chin et al., 2009). However, based on our results, it can realize that social influence did not affect behavioral intention to adopt Big Data, thus, regarding more investigation to support decision-makers is needed.

Theoretical implications, limitations, and future research

From the theoretical perspective, this study makes critical contributions to the field of logistics in SCM. First, by integrating the literature on BDA, SCM, and UTAUT, we validated a strong theoretical model. We adapted and applied a previously developed model for use with Brazilian SCM professionals, and the results validated it. The theoretical model explained 46% of behavioral intention, outperforming the 20% threshold in the literature (W. W. Chin, 1998; Martins et al., 2014). As previously mentioned, since our results regarding discriminant validity are consistent with those in the literature, they support our hypothesized structural paths. In other words, the model actually measures the behavioral intention to adopt Big Data by SCM professionals.

Our findings reveal that facilitating conditions are a good predictor of the behavioral intention to use Big Data. Future research could focus on an in-depth understanding of the enablers of facilitating conditions, as well as its barriers. In the proposed model, in line with a prior study (Alalwan et al., 2017), trust was a good predictor of performance expectancy and behavior intention. On the other hand, social influence was not found to be a good predictor of behavioral intention, following the results reported in Alalwan et al. (2017). This finding suggests the need for further investigation of the role of social influence in Big Data adoption and other technologies that are emerging in the SCM field.

This study suffers from some limitations. We believe that, first, a moderator variable could be incorporated into the model (Venkatesh et al., 2003, 2012) to capture the nuances and differences in the sample, such as industry, gender, and experience. Second, because of the scarcity of Brazilian studies on Big Data adoption, our findings cannot be compared with other similar works in this context. However, it opens up opportunities for scholars and practitioners to apply the validated model and to adapt it to other contexts. Third, the adopted model was tested in an emerging economy; because of this, the results cannot be generalized globally. Consequently, obtaining more empirical evidence by applying the adopted model in other countries could be an exciting stream for future research.

Finally, this study was one of the first attempt to understand the behavioral intention to adopt Big Data by Brazilian SCM professionals. There is an urgent need and opportunities for additional investigations on this and other cutting-edge technologies (e.g., blockchain, internet of things, 3D printing, etc.), regarding the relationship, as also compare the hypotheses of this model into other contexts.

CONCLUSION

The purpose of this study was to gain an in-depth understanding of Big Data behavior intention among Brazilian SCM professionals and to adjust and apply a model that captures the constructs of adoption behavior. In this regard, our study contributes to a more thorough understanding of the intention to adopt BDA in the Brazilian SCM field.

The contributions of this study are threefold. First, supported by a strong theoretical literature (Akter et al., 2016; Alalwan et al., 2017; Davis, 1989; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019) we adapted and applied a model to understand behavioral intention concerning Brazilian SCM professionals. Second, our findings provide strong implications for theory and practice. For instance, one implication is that only facilitating conditions, and trust were good predictors of behavioral intention. Contrary to findings of previous studies (Venkatesh et al., 2003, 2012), social influence was not a predictor of behavioral intention, but this result is in line with the recent findings reported by Alalwan et al. (2017). Third, both performance expectancy and effort expectancy were not good predictors of behavioral intention. This interesting finding opens up opportunities to further exploration of this behavior. Finally, our study contributes to fill a gap in the Brazilian empirical literature on Big Data in SCM, while simultaenously motivates logistics and SCM scholars to advance this stream of research.

REFERENCES

Abawajy, J. (2015). Comprehensive analysis of big data variety landscape. International journal of parallel. *Emergent* and *D*istributed *Systems, 30*(1), 5-14. doi:10.1080/17445760.2014.925548

Ahmad, S. Z., & Khalid, K. (2017). The adoption of M-government services from the user's perspectives: Empirical evidence from the United Arab Emirates. *International Journal of Information Management*, *37*(5), 367-379. doi:10.1016/j.ijinfomgt.2017.03.008

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, *182*, 113-131. doi:10.1016/j.ijpe.2016.08.018

Alalwan, A. A., Dwivedi, Y. K., & Rana, N. P. (2017). Factors influencing adoption of mobile banking by Jordanian bank customers: Extending UTAUT2 with trust. *International Journal of Information Management*, *37*(3), 99-110. doi:10.1016/j.ijinfomgt.2017.01.002

Autry, C. W., Williams, B. D., & Golicic, S. (2014). Relational and process multiplexity in vertical supply chain triads: An exploration in the U.S. restaurant industry. *Journal of Business Logistics*, *35*(1), 52-70. doi:10.1111/jbl.12034

Azjen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Batara, E., Nurmandi, A., Warsito, T., & Pribadi, U. (2017). Are government employees adopting local e-government transformation? *Transforming Government: People, Process and Policy*, *11*(4), 612–638. doi:10.1108/TG-09-2017-0056

Bibri, S. E., & Krogstie, J. (2017). ICT of the new wave of computing for sustainable urban forms: Their big data and context-aware augmented typologies and design concepts. *Sustainable Cities and Society*, *32*, 449-474. doi:10.1016/j.scs.2017.04.012

Carter, C. R., Rogers, D. S., & Choi, T. Y. (2015). Toward the theory of the supply chain. *Journal of Supply Chain Management*, *51*(2), 89-97. doi:10.1111/jscm.12073

Chin, A., Wafa, S., & Ooi, A.-Y. (2009). The effect of internet trust and social influence towards willingness to purchase online in Labuan, Malaysia. *International Business Research*, *2*(2), 72-81.

Chin, W. W. (1998). Issues and opinion on structural equation modeling. *MIS Quarterly*, *22*(1), 7-16.

Choi, T. Y., Dooley, K. J., & Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: Control versus emergence. *Journal of Operations Management*, *19*(3), 351-366. doi:10.1016/S0272-6963(00)00068-1

Colquitt, J. A., & Rodell, J. B. (2011). Justice, trust, and trustworthiness: A longitudinal analysis integrating three theoretical perspectives. *Academy of Management Journal*, *54*(6), 1183–1206. doi:10.5465/amj.2007.0572

Croll, A. (2015). *Data: Emerging trends and technologies*. Retrieved from http://www.oreilly.com/data/free/data-emerging-trends-and-technologies.csp

Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review*. Retrieved from https://hbr.org/

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319-340. doi:10.2307/249008

Domo. (2017). Data never sleeps 4.0. Retrieved from https://www.domo.com/blog/data-never-sleeps-4-0/

Dwivedi, Y. K., Rana, N. P., Janssen, M., Lal, B., Williams, M. D., & Clement, M. (2017). An empirical validation of a unified model of electronic government adoption (UMEGA). *Government Information Quarterly*, *34*(2), 211–230. doi:10.1016/j.giq.2017.03.001

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897-904. doi:10.1016/j.jbusres.2015.07.001

Farooq, M. S., Salam, M., Jaafar, N., Fayolle, A., Ayupp, K., Radovic-Markovic, M., & Sajid, A. (2017). Acceptance and use of lecture capture system (LCS) in executive business studies: Extending UTAUT2. *Interactive Technology and Smart Education*, *14*(4), 329-348. doi:10.1108/ITSE-06-2016-0015

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39-50. Doi:10.2307/3151312

Francisco, K., & Swanson, D. (2018). The supply chain has no clothes: Technology adoption of blockchain for supply chain transparency. *Logistics*, *2*(1), 2. doi:10.3390/logistics2010002

Gefen, Karahanna, & Straub. (2003). Trust and TAM in online shopping: An integrated model. *MIS Quarterly*, *27*(1), 51-90. doi:10.2307/30036519

Giannakis, M., & Louis, M. (2016). A multi-agent based system with big data processing for enhanced supply chain agility. *Journal of Enterprise Information Management*, *29*(5), 706-727. doi:10.1108/JEIM-06-2015-0050

Gobble, M. M. (2013). Big data: The next big thing in innovation. *Research-Technology Management*, *56*(1), 64-67. doi:10.5437/08956308X5601005

Grawe, S. J., Daugherty, p. J., & Ralston, p. M. (2015). Enhancing dyadic performance through boundary spanners and innovation: An assessment of service provider-customer relationships. *Journal of Business Logistics*, *36*(1), 88-101. doi:10.1111/jbl.12077

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*, *70*, 308-317. doi:10.1016/j.jbusres.2016.08.004

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049-1064. doi:10.1016/j.im.2016.07.004

Hair, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM) An emerging tool in business research. *European Business Review*, *26*(2), 106-121. doi:10.1108/EBR-10-2013-0128

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on Partial Least Squares Structural Equation Modeling*. 2nd ed.. Thousand Oaks, CA: Sage Publications.

Han, J. H., Wang, Y., & Naim, M. (2017). Reconceptualization of information technology flexibility for supply chain management: An empirical study. *International Journal of Production Economics*, *187*, 196-215. doi:10.1016/j.ijpe.2017.02.018

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, *154*, 72-80. doi:10.1016/j.ijpe.2014.04.018

Hazen, B. T., Skipper, J. B., Ezell, J. D., & Boone, C. A. (2016). Big data and predictive analytics for supply chain sustainability: A theory-driven research agenda. *Computers & Industrial Engineering*, *101*, 592-598. doi:10.1016/j.cie.2016.06.030

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. In R. Sinkovics & P. Ghauri (Eds.), *New challenges to international marketing* (Vol 20, pp. 277-319). Emerald Group Publishing Limited.

Huang, T. C. K., Liu, C. C., & Chang, D. C. (2012). An empirical investigation of factors influencing the adoption of data mining tools. *International Journal of Information Management*, *32*(3), 257-270. doi:10.1016/j.ijinfomgt.2011.11.006

Kache, F., & Seuring, S. (2017). Challenges and opportunities of digital information at the intersection of big data analytics and supply chain management. *International Journal of Operations & Production Management*, *37*(1), 10-36. doi:10.1108/IJOPM-02-2015-0078

Kune, R., Konugurthi, p. K., Agarwal, A., Chillarige, R. R., & Buyya, R. (2016). The anatomy of big data computing. *Software: Practice and Experience*, *46*(1), 79-105. doi:10.1002/spe.2374

Mamonov, S., & Benbunan-Fich, R. (2017). Exploring factors affecting social e-commerce service adoption: The case of Facebook gifts. *International Journal of Information Management*, *37*(6), 590-600. doi:10.1016/j.ijinfomgt.2017.05.005

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute*. Retrieved from https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation

Martins, C., Oliveira, T., & Popovič, A. (2014). Understanding the internet banking adoption: A unified theory of acceptance and use of technology and perceived risk application. *International Journal of Information Management*, *34*(1), 1–13. doi:10.1016/j.ijinfomgt.2013.06.002

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709-734. doi:10.2307/258792

Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). Defining supply chain management. *Journal of Business Logistics*, *22*(2), 1-25. doi:10.1002/j.2158-1592.2001.tb00001.x

Morris, M. G., & Venkatesh, V. (2000). Age differences in technology adoption decisions: Implications for a changing work force. *Personnel Psychology*, *53*(2), 375-403. doi:10.1111/j.1744-6570.2000.tb00206.x

Munshi, A. A., & Mohamed, Y. A. R. I. (2017). Big data framework for analytics in smart grids. *Electric Power Systems Research*, *151*, 369-380. doi:10.1016/j.epsr.2017.06.006

Nobre, G. C., & Tavares, E. (2017). Scientific literature analysis on big data and internet of things applications on circular economy: A bibliometric study. *Scientometrics*, *111*(1), 463-492. doi:10.1007/s11192-017-2281-6

Nunnally, J. C. (1978). *Psychometric Theory: 2d Ed*. New York, NY: McGraw-Hill.

Oliveira, T., Faria, M., Thomas, M. A., & Popovič, A. (2014). Extending the understanding of mobile banking adoption: When UTAUT meets TTF and ITM. *International Journal of Information Management*, *34*(5), 689-703. doi:10.1016/j.ijinfomgt.2014.06.004

Oliveira, T., Faria, M., Thomas, M. A. & Popovič, A. (2014). Extending the understanding of mobile banking adoption: When UTAUT meets TTF and ITM. *International Journal of Information Management, 34*(5), 689-703. doi:10.1016/j.ijinfomgt.2014.06.004

Phillips-Wren, G., & Hoskisson, A. (2015). An analytical journey towards big data. *Journal of Decision Systems*, *24*(1), 87-102. doi:10.1080/12460125.2015.994333

Queiroz, M. M., & Telles, R. (2018). Big data analytics in supply chain and logistics: An empirical approach. *International Journal of Logistics Management*, *29*(2), 767-783. doi:10.1108/IJLM-05-2017-0116

Queiroz, M. M., & Wamba, S. F. (2019). Blockchain adoption challenges in supply chain: An empirical investigation of the main drivers in India and the USA. *International Journal of Information Management*, *46*, 70-82. doi:10.1016/j.ijinfomgt.2018.11.021

Richey, R. G., Morgan, T. R., Lindsey-Hall, K., & Adams, F. G. (2016). A global exploration of big data in the supply chain. *International Journal of Physical Distribution & Logistics Management*, *46*(8), 710-739. doi:10.1108/IJPDLM-05-2016-0134

 CC BY

Riffai, M. M. M. A., Grant, K., & Edgar, D. (2012). Big TAM in Oman: Exploring the promise of on-line banking, its adoption by customers and the challenges of banking in Oman. *International Journal of Information Management*, *32*(3), 239-250. doi:10.1016/j.ijinfomgt.2011.11.007

Ringle, C. M., Wende, S., & Becker, J. M. (2015). SmartPLS 3. Boenningstedt, Germany: SmartPLS GmbH.

Rotella, p. (2012). Is data the new oil? *Forbes*. Retrieved from http://www.forbes.com/sites/perryrotella/2012/04/02/is-data-the-new-oil/#44ed5b8e77a9

Sabi, H. M., Uzoka, F. E., Langmia, K., & Njeh, F. N. (2016). Conceptualizing a model for adoption of cloud computing in education. *International Journal of Information Management*, *36*(2), 183-191. doi:10.1016/j.ijinfomgt.2015.11.010

Schoenherr, T., & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: Current state and future potential. *Journal of Business Logistics*, *36*(1), 120-132. doi:10.1111/jbl.12082

Shim, S., Lee, B., & Kim, S. L. (2018). Rival precedence and open platform adoption: An empirical analysis. *International Journal of Information Management*, *38*(1), 217-231. doi:10.1016/j.ijinfomgt.2017.10.001

Srinivasan, R., & Swink, M. (2018). An investigation of visibility and flexibility as complements to supply chain analytics: An organizational information processing theory perspective. *Production and Operations Management*, *27*(10), 1849-1867. doi:10.1111/poms.12746

Stock, J. R., & Boyer, S. L. (2009). Developing a consensus definition of supply chain management: A qualitative study. *International Journal of Physical Distribution & Logistics Management*, *39*(8), 690-711. doi:10.1108/09600030910996323

Strawn, G. (2012). Scientific research: How many paradigms? *EDUCAUSE Review*. Retrieved from http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=EJ970900

Sun, J., & Teng, J. T. C. (2017). The construct of information systems use benefits: Theoretical explication of its underlying dimensions and the development of a measurement scale. *International Journal of Information Management*, *37*(5), 400-416. doi:10.1016/j.ijinfomgt.2017.04.010

Venkatesh, V., & Brown, S. A. (2001). A longitudinal investigation of personal computers in homes: Adoption determinants and emerging challenges. *MIS Quarterly*, *25*(1), 71-102. doi:10.2307/3250959

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425-478. doi:10.2307/30036540

Venkatesh, V., Thong, J., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157-178. doi:10.2307/41410412

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How "big data" can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, *165*, 234-246. doi:10.1016/j.ijpe.2014.12.031

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. fan, Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356-365. doi:10.1016/j.jbusres.2016.08.009

Wamba, S. F. (2018, March). Social media use in the workspace: applying an extension of the technology acceptance model across multiple countries. In *World Conference on Information Systems and Technologies* (pp. 385-392). Springer, Cham.

Wang, C. S., Jeng, Y. L., & Huang, Y. M. (2017). What influences teachers to continue using cloud services? The role of facilitating conditions and social influence. *The Electronic Library*, *35*(3), 520-533. doi:10.1108/EL-02-2016-0046

Wang, G., Gunasekaran, A., Ngai, E. W. T., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, *176*, 98-110. doi:10.1016/j.ijpe.2016.03.014

Weerakkody, V., El-Haddadeh, R., Al-Sobhi, F., Shareef, M. A., & Dwivedi, Y. K. (2013). Examining the influence of intermediaries in facilitating e-government adoption: An empirical investigation. *International Journal of Information Management*, *33*(5), 716-725. doi:10.1016/j.ijinfomgt.2013.05.001

Wills, M. J. (2014). Decisions through data: Analytics in healthcare. *Journal of Healthcare Management*, *59*(4), 254-262.

Wu, K. J., Liao, C. J., Tseng, M. L., Lim, M. K., Hu, J., & Tan, K. (2017). Toward sustainability: Using big data to explore the decisive attributes of supply chain risks and uncertainties. *Journal of Cleaner Production*, *142*(Part 2), 663-676. doi:10.1016/j.jclepro.2016.04.040

Wu, K., Zhao, Y., Zhu, Q., Tan, X., & Zheng, H. (2011). A meta-analysis of the impact of trust on technology acceptance model: Investigation of moderating influence of subject and context type. *International Journal of Information Management*, *31*(6), 572-581. doi:10.1016/j.ijinfomgt.2011.03.004

Yadlapalli, A., Rahman, S., & Gunasekaran, A. (2018). Socially responsible governance mechanisms for manufacturing firms in apparel supply chains. *International Journal of Production Economics*, *196*, 135-149. doi:10.1016/j.ijpe.2017.11.016

Youngberg, E., Olsen, D., & Hauser, K. (2009). Determinants of professionally autonomous end user acceptance in an enterprise resource planning system environment. *International Journal of Information Management*, *29*(2), 138-144. doi:10.1016/j.ijinfomgt.2008.06.001

Zhao, R., Liu, Y., Zhang, N., & Huang, T. (2017). An optimization model for green supply chain management by using a big data analytic approach. *Journal of Cleaner Production*, *142*, 1085-1097.

Zhao, Y., Ni, Q., & Zhou, R. (2018). What factors influence the mobile health service adoption? A meta-analysis and the moderating role of age. *International Journal of Information Management*, *43*, 342-350. doi:10.1016/j.ijinfomgt.2017.08.006

Zhong, R. Y., Newman, S. T., Huang, G. Q., & Lan, S. (2016). Big data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives. *Computers & Industrial Engineering*, *101*, 572-591. doi:10.1016/j.cie.2016.07.013

Zinn, W., & Goldsby, T. J. (2017a). In search of research ideas? Call a professional. *Journal of Business Logistics*, *38*(1), 4-5. doi:10.1111/jbl.12160

Zinn, W., & Goldsby, T. J. (2017b). The role of academic research in supply chain practice: How much are we contributing? Journal of Business Logistics, 38(4), 236-237. doi:10.1111/jbl.12175

FORUM

Submitted 10.19.2018. Approved 07.19.2019
Evaluated through a double-blind review process. Guest Scientific Editors: Eduardo de Rezende Francisco, José Luiz Kugler, Soong Moon Kang, Ricardo Silva, and Peter Alexander Whigham
Original version
DOI: http://dx.doi.org/10.1590/S0034-759020190606

MEASURING ACCESSIBILITY: A BIG DATA PERSPECTIVE ON UBER SERVICE WAITING TIMES

Medindo a acessibilidade: Uma perspectiva de Big Data sobre os tempos de espera do serviço da Uber

Medición de accesibilidad: Una perspectiva de Big Data sobre los tiempos de espera del servicio de la Uber

ABSTRACT

This study aims to relate information about the waiting times of ride-sourcing services, with specific reference to Uber, using socioeconomic variables from São Paulo, Brazil. The intention is to explore the possibility of using this measure as an accessibility proxy. A database was created with the mean waiting time data per district, which was aggregated to a set of socioeconomic and transport infrastructure variables. From this database, a multiple linear regression model was built. In addition, the stepwise method selected the most significant variables. Moran's I test confirmed the spatial distribution pattern of the measures, motivating the use of a spatial autoregressive model. The results indicate that physical variables, such as area and population density, are important to explain this relation. However, the mileage of district bus lines and the non-white resident rate were also significant. Besides, the spatial component indicates a possible relation to accessibility.

KEYWORDS | Accessibility, Big Data, Uber, space statistic, urban disparity.

RESUMO

O presente artigo busca relacionar informações sobre o tempo de espera de serviços de aluguel de carro, especificamente Uber, com variáveis socioeconômicas da cidade de São Paulo com a intenção de explorar a possibilidade uso dessas medidas como um proxy de acessibilidade. Foi montada uma base com a média dos dados de tempo de espera do serviço por distrito, que foi agregada a um conjunto de variáveis socioeconômicas e de infraestrutura de transporte. A partir dessa base foram elaborados modelos de regressão linear multipla (RLM), e utilizando o método stepwise foram selecionadas as variáveis mais significativas do modelo. Foi verificado padrão espacial das variáveis através do teste I de Moran, que motivou a elaboração de um modelo espacial autoregressivo (SAR). Os resultados indicam que variáveis físicas são importantes para essa relação, como área e densidade populacional, mas a quilometragem de linhas de ônibus no distrito a taxa de residentes não brancos, além do componente espacial, indica uma possível relação com acessibilidade.

PALAVRAS-CHAVE | *Acessibilidade, Big Data, Uber, estatística espacial, disparidade urbana.*

RESUMEN

El presente artículo busca relacionar informaciones sobre el tiempo de espera de servicios de alquiler de coches, específicamente Uber, con variables socioeconómicas de la ciudad de São Paulo con la intención de explorar la posibilidad de utilizar esas medidas como un proxy de accesibilidad. Se ha montado una base con la media de los datos de tiempo de espera del servicio por distrito, que se ha agregado a un conjunto de variables socioeconómicas y de infraestructura de transporte. A partir de esta base se elaboraron modelos de regresión lineal MLR, y utilizando el método stepwise se seleccionaron las variables más significativas del modelo. Se verificó el patrón espacial de las variables a través de la prueba I de Moran, que motivó la elaboración de un modelo espacial autoregresivo (SAR). Los resultados indican que las variables físicas son importantes para esa relación, como el área y la densidad de población, pero el kilometraje de líneas de autobús en el distrito, la tasa de residentes no blancos, además del componente espacial, indica una posible relación con accesibilidad.

PALABRAS CLAVE | *Accesibilidad, Big Data, Uber, estadística espacial, disparidad urbana.*

ANDRÉ INSARDI[1]
andre.insardi@espm.br
ORCID: 0000-0003-3782-3505

RODOLFO OLIVEIRA LORENZO[2]
rodolfolo@uol.com.br
ORCID: 0000-0003-4847-9201

[1]Escola Superior de Propaganda e Marketing, São Paulo, SP, Brazil

[2]Fundação Getulio Vargas, Escola de Administração de Empresas de São Paulo, São Paulo, SP, Brazil

 CC BY

INTRODUCTION

The introduction of ride-sourcing companies in the private urban mobility market has changed the habits of many urban inhabitants considerably, including its traditional users, private car owners, and commuters alike.

The increase in competition in a traditionally highly regulated market led to conflicts in many cities between the new services, the former suppliers of this market, taxis, and the local authorities. The difficulties faced by local authorities in framing these new services within existing legal frameworks is clearly illustrated in the case of San Francisco (Flores & Rayle, 2017).

In this context, ride-sourcing companies tried to build an environmentally friendly image, selling themselves as ride-sharing services (Flores & Rayle, 2017). Beyond that, they claimed the capacity to reduce the number of vehicles on the street (and their emissions) and to offer better and cheaper services to areas formerly neglected by the taxis.

In this debate, a number of studies tried to shed light on the effects of the resulting transformations. Jin, Kong, Wu, and Sui (2018) review recent literature regarding the effects of the ride-sourcing services. Among the matters discussed are the economic efficiency and social equity of these services. Broadly, there seems to be evidence to support the increase in private transportation services in peripheral and low-income areas formerly neglected by taxis. Despite that, the customer profile of the service is younger, richer, and better educated than the mean individual of the population; the article points to a possible exclusion from the service along the lines of a "digital divide," in terms of both generation and income (Jin et al., 2018).

However, another essential aspect of ride-sourcing services is their capacity to generate high quality "Big Data" that can be used to evaluate several questions about the service. The access to this data allows for rich analysis with great detail and can further reveal the interplay between other modes of transportation and ride-sourcing (Jin et al., 2018).

In particular, the possibility of analyzing urban accessibility with this data source is interesting. As an essential component of accessibility, the transport network distribution (Páez, Scott, & Morency, 2012) is a recurrent research subject among geographers, urban planners, and social scientists. Recently, the methods and approach of this field of studies are being transformed. This is a result of the influence of, among other things, Big Data and opening dialogs with other disciplines (Schwanen, 2016). Simultaneously, the field continues to make itself relevant: whilst distance friction is a reality, accessibility will continue to be a useful concept to describe urban experience (Páez et al., 2012).

Particularly considering Big Data approaches, Letouzé and Jütting (2015) affirm that the official bodies responsible for the production of official statistics and indicators, including academic bodies, must be aware of the evolutions in "Big Data." This is necessary both to profit from new tools and approaches, together with the scientific rigor of validation and analysis, as well as to face the world of Big Data as an inestimable source of data for the advance of scientific research. The potential of Big Data tools is also of considerable relevance to public managers and policy makers (Kim, Trimi, & Chung, 2014). In a multimodal transport network, the usage of different Big Data tools can help to better regulate the private transport supply and to better deliver public transportation according to user needs (Kim et al., 2014; Lessa, Lobo, & Cardoso, 2019).

In this context, the growth of global tech companies in mobility, such as Uber, Cabify and Lyft, has made them considerable players in this field. For instance, Uber has a daily average of 15 million rides across the world. In Brazil, the company provides its service in 100 cities with a network of 500 000 drivers and more than 20 million users.

Uber has developed, commercialized, and operates the application for smartphones that allows consumers to request rides from partnered drivers. In the process of requesting the rides, the Uber tool provides estimates of waiting time and travel cost on the user's app and in the web environment through public application programming interfaces (APIs).

These APIs generate estimates through the analysis of ride history in the user's region and the supply-and-demand curve of Uber's cars (Cohen et al., 2016) as can be verified in the available documentation (https://developer.uber.com).

Wang and Mu (2018) and Hughes and Mackenzie (2016) propose the usage of these estimates as a possible measure for accessibility. The interferences of these new services in the transportation environment and the easy access to the tools Uber provides have already motivated some studies (Hall & Krueger, 2016; Hughes & MacKenzie, 2016; Wang & Mu, 2018; Zhou, Wang, & Li, 2017). This opens the opportunity for empirical exploration of Uber's fleet in light of accessibility theory.

This study's aim is to use the Big Data tools developed by Uber, one of the largest ride-sourcing providers in São Paulo, to generate data to conduct an exploratory study of a potential accessibility measurement.

As Uber's pricing algorithm follows the balance between supply and demand in order to influence driver's behavior (Hall, Horton, & Knoepfle, 2019), we assume that waiting times for Uber rides can reflect regional imbalances in the supply of cars. Following this line of thought, we have explored the relationship

 CC BY © RAE | São Paulo | **59(6)** | **November-December 2019** | 402-414 ISSN 0034-7590; eISSN 2178-938X

of waiting times with other variables associated with accessibility in similar literature (Hughes & MacKenzie, 2016; Lessa et al., 2019; Wang & Mu, 2018). Particularly, related to mobility and socioeconomic factors.

The results found are in contrast to previous studies (Hughes & MacKenzie, 2016; Wang & Mu, 2018), regarding the relevance of variables in the context of São Paulo. Keeping in mind the warning in Schwanen (2016) about generalizing conclusions with Big Data to different local contexts, further studies can compare different cities to elucidate how the local context can be better taken into account in similar approaches. Moreover, further validation of this data source can build a new tool capable of facilitating the decision making of transport planners and public policy managers in cities (Kim et al., 2014; Letouzé & Jütting, 2015).

LITERATURE REVISION AND RESEARCH QUESTION

This section begins with the conceptualization of "Big Data" and it's positioning in the current technology market, and the conceptualizing Uber's estimation tool in light of the "Big Data" concept. It is followed by a revision of the concept of accessibility, its normative and positivist dispositions, and its managerial context.

Big Data

A classical definition of the Big Data movement takes into consideration the features of the data produced in the virtual environments of massive user presence: Volume, Variety, and Velocity, the three Vs (McAfee & Brynjolfsson 2012).

According to this definition, the data generated by the new ways of using technology and applications generates relatively large databases, arriving at the scale of Petabytes or Exabytes. The various forms of data generation (such as photo posts, comments on social networks, reactions to comments from others, videos and audios, etc.) are responsible for generating heterogeneous databases in contrast to structured databases. The timing of the generation of these databases is almost instantaneous, demanding real time processing in some cases. These aspects inform the techniques capable of handling the data in this valuable timeframe.

This is not the only conceptual approach to Big Data. A more sociological view tries to describe the movement with three "Cs": Crumbs, Capacities, and Communities (Letouzé & Jütting, 2015).

"Crumbs" is a reference to the nature of the data collected in relation to the behavior of users in these new applications. These users leave behind traces of their activities while interacting and these traces, or crumbs, constitute the databases to be analyzed in Big Data. "Capacities" are the techniques, both statistical and programming, used to manipulate these data and extract information. The third concept, "Communities," refers to the behavior patterns of the producers of Big Data environments inside specific communities whose members share ideas with specific language and common validation methods. These communities can be established in open and collaborative environments, such as OpenSource communities, or in more restrictive ones, like tech groups in big corporations with access to big databases (Letouzé & Jütting, 2015).

The Uber API in the Big Data context

Uber's estimation tool used in this article is by both definitions a Big Data tool. In relation to the three Vs, a huge volume of rides results in data constantly feeding the tool's algorithm with high speed interpretations of spatial non-structured data (Cohen et al., 2016).

At the same time, the data that feeds the algorithm are traces of drivers' and users' activities, or crumbs. The tools that process the data received online at this high rate are also tools included in the term "Capacities" developed by Big Data environments. In addition, the community of operators in the system of transport startups, particularly Uber, fits the "Communities" concept.

Furthermore, Pääkkönen and Pakkala's (2015) description of Big Data software architecture suits Uber's API architecture and technology as described by its documentation (https://developer.uber.com) very well.

Accessibility

Discussing individuals' access to urban mobility in the city is an important subject. There are a number of different approaches, from the use of traditional statistical surveys (Metrô, 2008) to computational simulations of urban mobility behavior in a given urban territory (Krajzewicz, Erdmann, Behrisch, & Bieker, 2012), including approaches close to the present one, using social media data to trace mobility behavior (Noulas, Scellato, Lambiotte, Pontil, & Mascolo, 2012).

 CC BY ISSN 0034-7590; eISSN 2178-938X

In big urban centers, comparable in size and regional importance to São Paulo, knowing people's mobility is essential to understanding urban dynamics as the sheer size of the city feeds its capacity to attract more people. (Aranha, 2005). This has implications for mobility solutions for the locomotion demand. One of the main factors for measuring individuals' capacity to move inside the city is their routine travel times.

In São Paulo, this variable reflects socioeconomic contexts, indicating that access to mobility is differentially distributed in terms of space and social contexts (Morandi et al., 2016). However, mobility as a concept of the capacity to move in the city is limited in the sense of not considering the conditions of locomotion.

A more comprehensive concept in this sense is accessibility (Litman, 2003; Páez et al., 2012; Stelder, 2016). It is possible to define accessibility as the potential to access spatially distributed opportunities (Wang & Mu, 2018). Hansen (1959) defines accessibility by the manner in which people interact with places. To illustrate, the amount of roads measured in kilometers in a determined region is a usual measure. Similarly in geography, accessibility is the measure of how a person participates in a determined activity (M. Kwan, 1998; Weibull, 1980).

Accessibility reflects spatial development that consists of transport network and distribution of opportunities, materialized in soil uses and occupations. A possible interpretation for this is as a temporal measure (time to access) (Lessa et al., 2019; Páez et al., 2012). As a practical example, it is possible to consider the travel time to work as a comparative measure to understand the balance in job occupation and the racial, economic, and gender disparities contained in urban area distributions (Preston & Mclafferty, 1999; Tribby & Zandbergen, 2012).

Geographers and social scientists have critically analyzed the economy and the inequality in transport, and its correlations with socioeconomic inequalities. Schwanen (2016) argues that transport distribution has sociospatial polarization intensified under capitalism dynamics because the transport infrastructure is an asset that attracts capital and investments, bringing job opportunities, more efficiency, and competition.

In discussing accessibility, Páez et al. (2012) define two epistemic approaches to studying the concept: the first one is a normative approach defined in terms of which accessibility parameters are to be considered as reasonable, or in other words, how much is reasonable for a person to travel. The second approach is positivism and is defined in terms of observed accessibility parameters, or how much people do travel. The normative approach analyzes travel expectations while positivism bases itself on the actual travel experience. This study will take into consideration the positivist approach as the data refers to actual Uber waiting times in São Paulo.

Studies such as that of Hughes and MacKenzie (2016), Lessa et al. (2019) and Wang and Mu (2018) address this positivist approach. In a more traditional fashion, Lessa et al. (2019) investigate the relations between travel times collected in origin-destination surveys and transit data and the distribution of public transport network infrastructure. Hughes and MacKenzie's (2016) approach to accessibility uses waiting times for Uber services and socioeconomic data in order to explore the spatial correlations between the dispersion of Uber's service and Seattle's socio-economic disparities. It can be measured by variables such as population density, average income per capita and percentage of non-whites. Wang and Mu (2018) do a cross-sectional study creating a spatial lag regression model that tests the relation of Uber waiting time with socioeconomic and transport infrastructure variables of the city of Atlanta.

As Hughes and MacKenzie (2016) and Wang and Mu (2018) point out, these preliminary studies are subject to economic and cultural influences from the researched location. In addition Letouzé and Jütting (2015), Schwanen (2016), and Kwan (2016) warn us about the possible existence of bias in databases from Big Data tools. This is extremely relevant to the discussion of the accessibility and use of databases of Big Data in the investigation of Uber waiting time as a proxy for accessibility in São Paulo, because the particular economic and cultural context is relevant for the outcomes of the analysis.

Inspired by these studies, this article discusses the following hypotheses in the context of São Paulo:

> H_1: The estimated waiting time when requesting a ride on Uber's platform can be used as a proxy for accessibility;
>
> H_2: Uber waiting time distribution relates to socioeconomic indicators' polarization.

METHODOLOGY

This section begins with a description of the data collected of estimated Uber waiting times and socioeconomic data followed by a descriptive analysis to define the data clippings for final analysis (Wang & Mu, 2018). Then, the construction of the multiple linear regression (MLR) models with stepwise method is explained. As the significant variables selected by this method were found by Moran's I test to have high spatial dependency, a final spatial autoregressive (SAR) model was calculated, as suggested by Wang & Mu (2018).

 CC BY ISSN 0034-7590; eISSN 2178-938X

Data and variables

To verify the hypotheses elaborated above, two sets of data were collected. The first is the data from Uber waiting times, and the second, the spatialized socioeconomic data from São Paulo.

Uber Data

The data were collected from the APIs present in Uber's Developers portal, accessible from the company's site (https://developer.uber.com), which cover the estimated waiting times of all Uber products during August 2018.

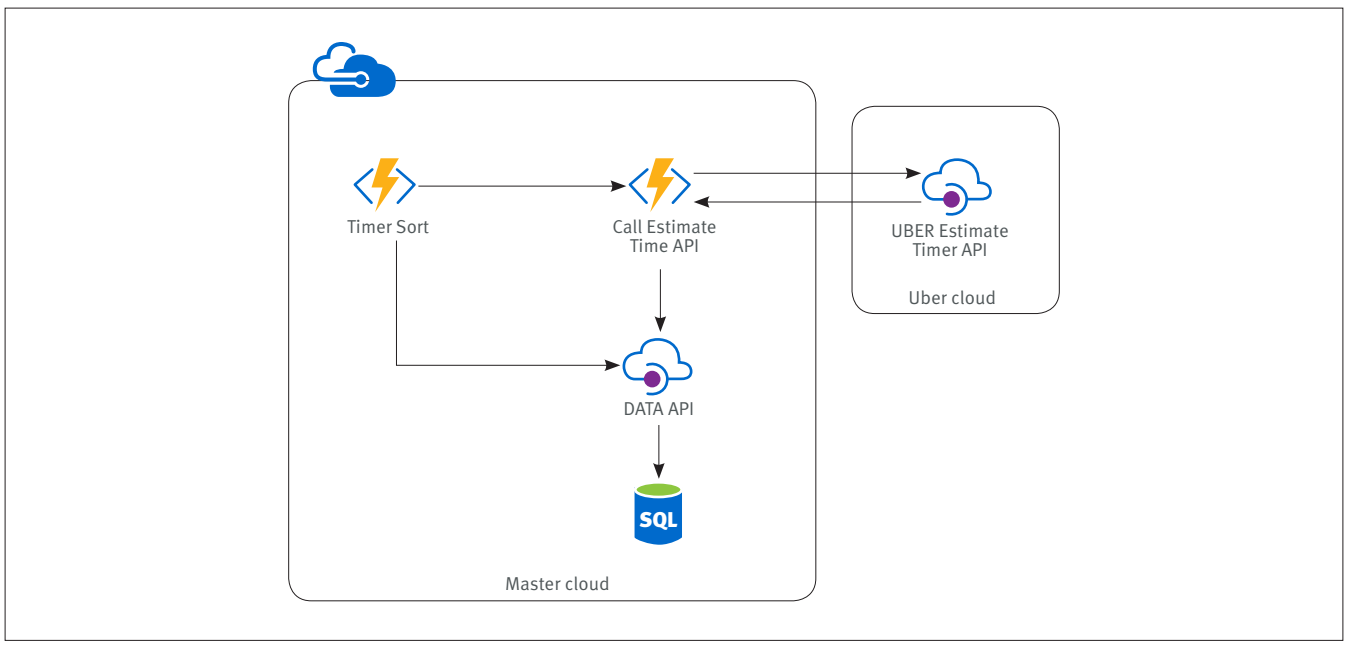
The city of São Paulo is divided into 96 districts, a territorial and administrative division that provides the local administration a certain degree of autonomy (Francisco, 2010). Following the methodology proposed by Wang and Mu (2018), the districts were used as the spatial unit of analysis. To guarantee that each district had at least one random sample point, the software developed for the data collection followed the logic below:

- Section the city in squares of 1km², amounting to 1720 squares;
- Randomly sample a point in each square;
- The sampled coordinates are used to consult the Uber waiting times API;
- If the return is successful, the result is stored;
- If an error flag is raised, three more attempts with the same coordinate are made before the attempt is stored as a null.

This process was repeated each 30 minutes during August 2018. More than 2,528,400 calls to the API were stored being a little more than 2,240,000 valid calls.

To process the data collected, a C# software program was developed following the concepts of cloud computing to access the Uber's APIs. This program was allocated in the Azure cloud computing service provided by Microsoft. Its architecture follows Figure 1.

Figure 1. **Diagram of the Uber data collection application**



Socioeconomic data

The choice of socioeconomic indicators was based on similar studies (Hughes & MacKenzie, 2016; Wang & Mu, 2018). The analysis units chosen were the districts of São Paulo. Comparable variables were found for population density, employment density, minority rate, mean income per capita, motorization rate (for cars and motorcycles), public transport infrastructure, and mean travel time to work. The socioeconomic data at the district level was collected from the 2010 IBGE (Brazilian Institute of Geography and Statistics) demographic census, the National Ministry of Work and Employment (RAIS – Anual Report of Social information), the SEADE's São Paulo State municipalities Indicators web portal, and the São Paulo Municipal portal of Geo-referenced information (http://geosampa.prefeitura.sp.gov.br) and is summarized in Exhibit 1.

CC BY
© RAE | São Paulo | **59(6)** | **November-December 2019** | 402-414 ISSN 0034-7590; eISSN 2178-938X

Exhibit 1. **São Paulo's socioeconomic data**

District data	Source	Date
Area (km²)	Data from IBGE. Collected from the portal "Indicadores dos Municípios Paulistas" (IMP) – SEADE Foundation	2009
Population	Original data from IBGE's demographic census annually readjusted by SEADE Foundation. Collected from the portal "Indicadores dos Municípios Paulistas" (IMP) – SEADE Foundation	2010/2018
Population Density	Computed from area and population data	2010/2018
Permanent particular households	Original data from IBGE's demographic census annually readjusted by SEADE Foundation. Collected from the portal "Indicadores dos Municípios Paulistas" (IMP) – SEADE Foundation	2010/2018
Income per Capita - Demographic Census (In current reals)	Original data from IBGE's demographic census. Collected from the portal "Indicadores dos Municípios Paulistas" (IMP) – SEADE Foundation	2010
Jobs (Commerce, Services, Transformation Industry, Civil Construction)	São Paulo's municipal portal "Infocidade." Original data source: Ministry of Work and Employment - Annual Report of Social Information (RAIS)	2010/2016
Employers (Commerce, Services, Transformation Industry, Civil Construction)	São Paulo's municipal portal "Infocidade." Original data source: Ministry of Work and Employment - Annual Report of Social Information (RAIS)	2010/2016
% of non-whites (Black, Pardos, Indigenous)	IBGE's 2010 Demographic Census	2010
Travel times	IBGE's Demographic Census Sample. Proportion of people, weighted, in each of the survey's class of travel time, by district.	2010
Household motorization rate	IBGE's Demographic Census Sample. Motorization rate by district (cars and motorcycles) from the weighted households of the Sample	2010
Bus Stops	Geosampa Portal	2018
Extension of Bus Lines (Km)	Geosampa Portal	2018
Metro Stations	Geosampa Portal	2018

The decennial Demographic census is one of Brazil's most important statistical products. The 2010 census presents two sets of data: the universal, which ideally comprehends every Brazilian, and the sample, in which the respondents (a statistical fraction of the populations) are asked a more detailed survey. The sample fraction varies between municipalities: in São Paulo around 5% of the households were included in the Sample. Across the whole of Brazil 10,7% were selected, or 6,192,332 households (IBGE, 2010).

The RAIS is the Ministry of Work and Employment's tool for management of Brazilian work relations. The data is compiled from statements made by businesses about the working situation between them and their employees. The declarations are mandatory for a series of businesses nominated by law (Ministério do Trabalho, 2016).

The SEADE Foundation is a nationally recognized statistical institution. The foundation is known for its technical capacity. It

 CC BY ISSN 0034-7590; eISSN 2178-938X

is also responsible, as part of the National Statistic System for producing data, for aggregating existing data in the interests of São Paulo State and its municipalities (Francisco, 2010).

São Paulo's municipal online platform GEOSAMPA provides a series of geo-referenced data on a range of issues, including the distribution of transportation infrastructure.

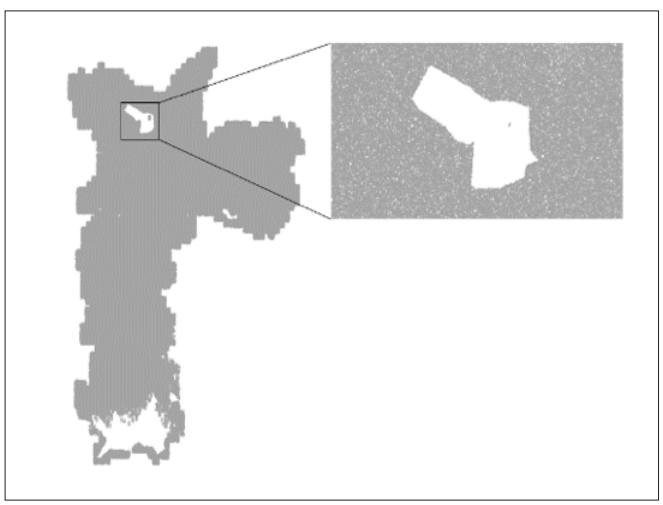
All data was searched or aggregated to the district level (Francisco, 2010; Hughes & MacKenzie, 2016; Lessa et al., 2019; Wang & Mu, 2018). The 2010 census Sample was used to account for the percentage of the population of each district that is in different classes of travel time to work/school, and also for the percentage of households in each district that possess a car and a motorcycle.

The socioeconomic data distribution follows a clear center periphery pattern with minor variations, and is associated with the distribution of infrastructure and public policies (Francisco, 2010; Torres, Marques, Ferreira, & Bitar, 2003). The relations of this pattern with the distribution of Uber waiting times can shed some light on the relation of accessibility and sociospatial composition, as attempted by Wang and Mu (2018).

Descriptive data analysis

The exploration of the 2,528,400 API calls showed two patterns: First, the general lack of estimates for some of Uber's products, and second, the complete absence of calls in some regions of the city, as is shown in Figure 2. It is possible to note the absence in the extreme south of the city and in a strip to the north, besides some spots spread over the city.

Figure 2. **Coordinate points for Uber waiting time estimates**



The identified causes were the presence of a mangrove area in the south that does not have road access. The northern strip and the other spots over the city are regions that Uber classifies as risk zones and it does not provide services there.

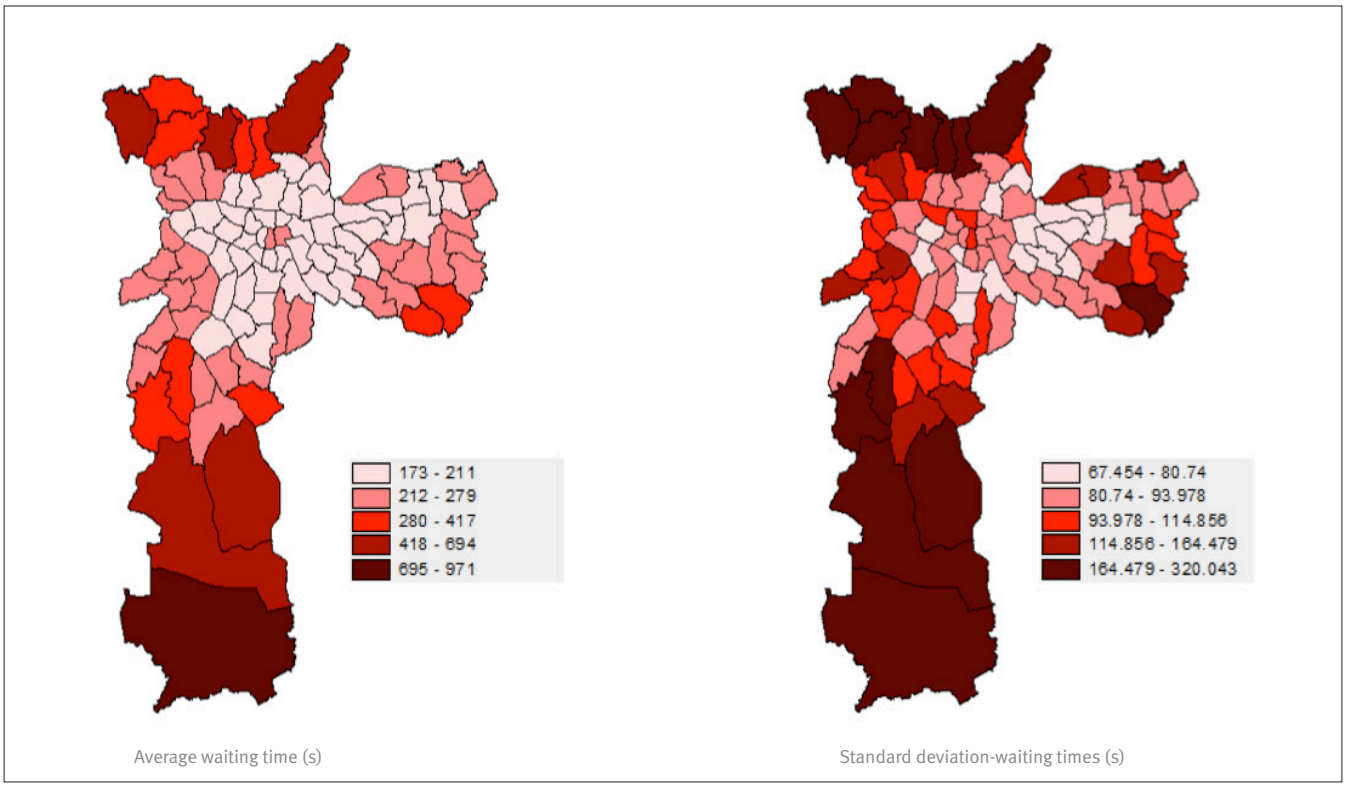
The cover of each service and its availability can be seen in Table 1. UberX, the most popular service Uber provides, has the largest coverage with almost 100% of time estimate responses registered. Consequently, it is the service that better reflects Uber's fleet spatial distribution. Therefore, UberX waiting times were adopted for the purpose of measuring accessibility.

Table 1. **Returns of calls of Uber waiting time estimates, by product**

Service	Number of successful calls	Average of estimated waiting time (seconds)	Standard deviation of estimated waiting time (seconds)
Bag	1,327,710	400	187.2
Bike Rack	58,758	598	280.6
Black	987,605	491	215.5
Black Bag	549,919	506	220.6
Pool	361,166	200	87.9
Select	1,713,733	340	170.6
Uber X	2,233,720	377	277.9

Filtering only data from the UberX product, a great amplitude in the averages and the standard deviations between districts can be seen in Figure 3. It is reasonable to consider that a district with a bigger fleet and more access shows less variation in waiting times. We could say more accessible districts show lower standard deviations (Wang & Mu, 2018). For this reason, two MLR models were created. The first having the average waiting time as a dependent variable, and the second having the standard deviation as the dependent variable.

Figure 3. **Comparison between the average and standard deviation of UberX estimate waiting times**



Average waiting time (s)

Standard deviation-waiting times (s)

Multiple linear regression (MLR)

From the principle of inference statistics, it is possible to make statements about the characteristics of a population with a sample of it. Regression analysis is the term that describes a family of methods that permits exploring and inferring the relation between two or more variables (Francisco, 2010; Hair, 2006).

For the construction of the MLR models, only the data from UberX aggregated by districts was used. The dependent variable for model 1 was the average waiting time and that for model 2 was the standard deviation. The independent variables used were:

- Area (Km²)
- Population
- Population density
- Income per capita – Demographic Census (in reals)
- Jobs (Commerce, Services, Transformation Industry, Civil Construction)
- Employers (Commerce, Services, Transformation Industry, Civil Construction)
- Proportion of non-white residents (Black, Pardos, and Indigenous)
- Travel time
- Rate of household car and motorcycle motorization
- Number of bus stops
- Bus line length
- Quantity of bus lines
- Number of metro stations

The software R and its extensions "stats" and "car" were used for the computation of the regressive models.

Spatial auto regressive model (SAR)

Francisco (2010) suggests that before creating an SAR model it is convenient to verify the spatial auto-correlation of the dependent variable. The literature uses a measure established by Moran. The Moran index is an indicator of the correlation between the value of the observed variable in a spatial unit of analysis and the values of that variable on the unit's region (its neighbors).

After the verification of the geographic auto-correlation of Uber waiting times through Moran's I, the highly significant variables of the MLR model were selected and the SAR was calculated. Francisco (2010) defines SAR as a regression model capable of incorporating the spatial neighbors' matrix (or spatial proximity) as a part of the explanatory variables.

The GeoDa software version 1.12 was used to build the SAR model.

CC BY
© RAE | São Paulo | **59(6)** | **November-December 2019** | 402-414 ISSN 0034-7590; eISSN 2178-938X

RESULTS

Table 2 summarizes the results of the MLR models with the dependent variable as the average of UberX waiting times and standard deviation of waiting times. It is possible to note that the average time model has a better degree of explanation as it has an R^2 of 0.893 against an R^2 of 0.717 for the standard deviation model.

Table 2. **Average and standard deviation regressions result for UberX**

DV		Average Uber X			Std. Deviation Uber X		
		Coefficient	Std. Error	p-value	Coefficient	Std. Error	p-value
(Intercept)		943.00	750.40	0.213	39.35	515.30	0.939
QTLINBUS2018	Bus lines' quantity	0.13	0.19	0.514	0.12	0.13	0.353
KMLINBUS2018	Bus lines' length (Km)	-0.12	0.07	0.080	-0.09	0.05	0.057
QTPONTBUS2018	Number of bus stops	-0.12	0.13	0.361	0.01	0.09	0.902
QTESTMETRO2018	Quantity of Metro stations	0.91	5.22	0.862	-1.14	3.58	0.751
RENDP2010	Income per Capita	-0.00	0.01	0.922	0.00	0.01	0.515
ARE1	Area (km^2)	3.55	0.31	0.00	0.74	0.21	0.001
POP2018	Population	0.00	0.00	0.492	-0.00	0.00	0.894
DENPOP2018	Population density	-0.00	0.00	0.007	-0.01	0.00	0.000
DOMP2018	Number of particular permanent households	-0.00	0.00	0.646	0.00	0.00	0.546
ESTAB2016	Employers	-0.00	0.01	0.970	-0.00	0.00	0.791
EMP2016	Jobs	0.00	0.00	0.384	0.00	0.00	0.547
PNBRAN2010	Proportion of non-whites	300.20	77.69	0.000	202.20	53.35	0.000
TEMP2010_5MIN	Travel time - up to 5 minutes	-587.50	1,166.00	0.616	111.00	800.50	0.890
TEMP2010_30MIN	Travel time - from 6 to 30 minutes	-978.90	785.70	0.217	37.68	539.50	0.945
TEMP2010_60MIN	Travel time - from 31 to 60 minutes	-278.40	831.70	0.739	180.80	571.10	0.752
TEMP2010_120MIN	Travel time - from 61 to 120 minutes	-1,046.	805.20	0.198	-141.10	552.90	0.799
TEMP2010_121MIN	Travel time - more than 121 minutes	-1,443.	992.80	0.150	-215.90	681.80	0.752
TEMP2010_0MIN	No travel	-821.30	765.00	0.286	10.86	525.30	0.984
PDOMC2010	Car - rate of household motorization	-139.10	367.70	0.706	-164.10	252.50	0.518
PDOMM2010	Motorcycle - rate of household motorization	37.73	82.42	0.648	25.17	56.60	0.658

A stepwise MLR model was created analyzing the collinearity of the independent variables of the final model, it was considered convenient to work with variance inflation factors (VIF's) of less than 5 as recommended by Batterham, Tolfrey, and George (1997).

Besides that, the model selected was the mean of waiting time, because of its greater power of explanation. Table 3 shows the result of the model. Meanwhile, Table 4 shows the collinearity analysis of the selected independent variables for this final model. We can observe a high degree of significance for bus line length, district area, population density, jobs, percentage of non-whites, and travel times over 120 minutes, with an R-squared of 0.879. These elements corroborate the positivist view of Lessa et al.'s (2019) and Páez et al.'s (2012) studies.

Table 3. **Stepwise regression results for UberX average waiting times**

DV		Average Uber X		
		Coefficient	Std. Error	p-value
(Intercept)		171.00	18.17	0.000
KMLINBUS2018	Bus lines' length (Km)	- 0.0786	0.02	0.002
ARE1	Area (km²)	3.634	0.23	0.000
DENPOP2018	Population density	- 0.0035	0.0009	0.000
EMP2016	Jobs	0.00021	0.0001	0.097
PNBRAN2010	Proportion of non-whites	295.70	44.57	0.000
TEMP2010_121 MIN	Travel time - more than 121 minutes	- 1,239,00	422.50	0.004

Table 4. **Variance inflation factors analysis**

VIF		Average Uber X
		Coefficient
KMLINBUS2018	Bus line length (Km)	1.181254
ARE1	Area (km²)	2.139971
DENPOP2018	Population density	1.317234
EMP2016	Jobs	1.894758
PNBRAN2010	Proportion of non-whites	2.462318
TEMP2010_121MIN	Travel time - more than 121 minutes	2.615456

Figure 4 shows the geographical dependence of the average Uber X waiting time, with a Moran's I of 0.59, which shows the geographic slope of the dependent variable. In Figure 5, we note the distribution of the variable and its geographical dependence through the neighborhood as we can notice uniform areas with below time and low attendance.

Figure 4. **Uber X average waiting time Moran's I**

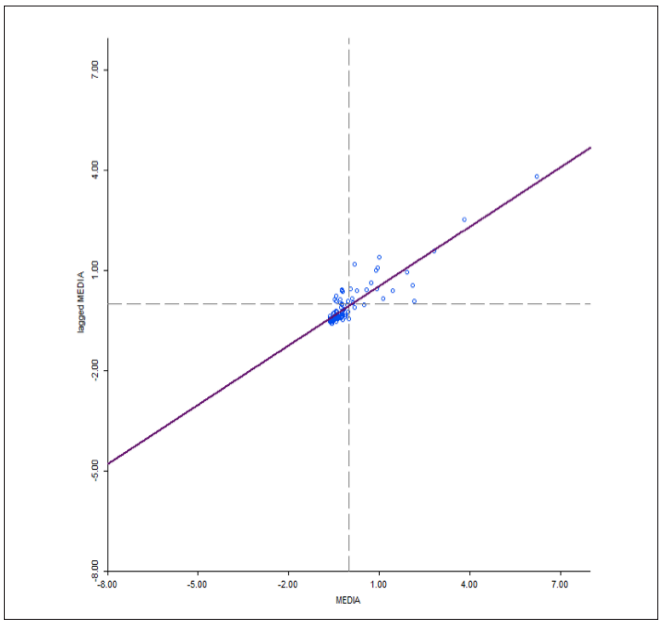
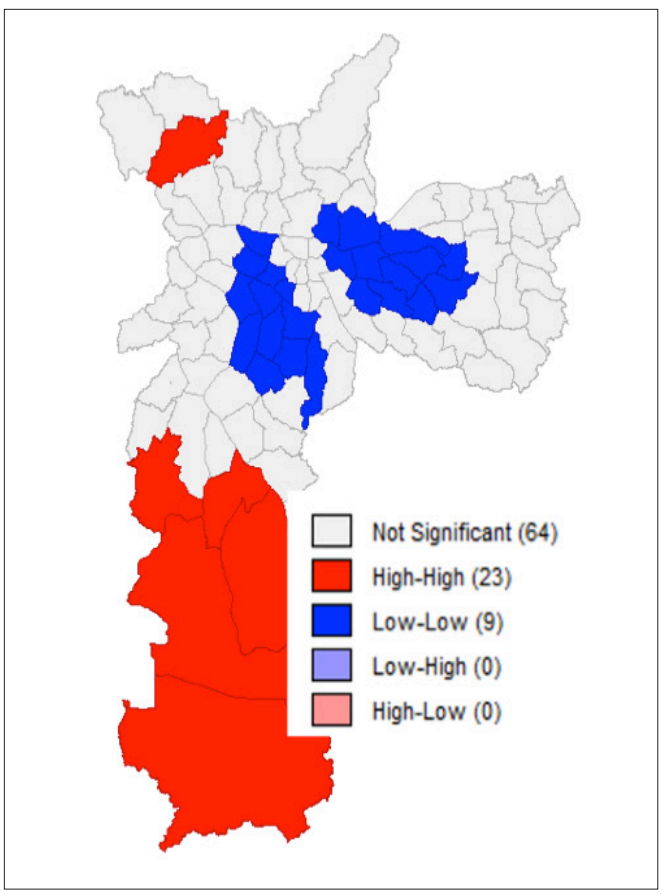


Figure 5. **Uber X average waiting time for neighborhood clusters**



From the variables in Table 3 we generated a spatial regression model for the average Uber X waiting time. The results presented in Table 5 show a high degree of significance for the variables bus line length, district area, population density, and percentage of non-whites. With the incorporation of the geographic factor into the model, the R-squared jumps to 0.89.

Table 5. **Spatial lag regression results of Uber X**

DV		Average Uber X		
		Coefficient	Std. Error	p-value
W_MEDIA		0.406	0.092	0.000
CONSTANT		114.573	23.246	0.000
KMLINBUS2018	Bus lines' length (Km)	-0.073	0.022	0.001
ARE1	Area (km²)	2.455	0.263	0.001
DENPOP2018	Population density	-0.002	0.001	0.005
PNBRAN2010	Proportion of non-whites	211.038	44.297	0.000

Observing the coefficients of the final SAR model, we can say that the number of bus lines (-0.073) and the population density (-0.0023) negatively influence Uber waiting time, as is expected according to Wang and Mu (2018). The logic is that the highest concentration of population and its transport flow would attract the Uber drivers to these regions, boosting supply and thus providing a shorter waiting time for service.

By contrast, it is worth noting the coefficient of the nonwhite percentage (211.03) variable, which shows an increase in waiting time when the concentration of minorities is identified in the region. The possibility of this relationship—worsening travel and waiting times versus minority distribution—has been approached on a recurring basis in the literature on accessibility (Flores & Rayle, 2017; Hughes & MacKenzie, 2016; Páez et al., 2012; Wang & Mu, 2018). As suggested by Wang and Mu (2018), the economic cultural differences between São Paulo and Atlanta may be the reason for the discrepancy between the results obtained. This occurrence motivates us to continue the debate on the regional idiosyncrasies of accessibility.

When analyzing the results, we cannot confirm H_1, that is, we cannot affirm that the estimation of service waiting time can be used as an accessibility proxy. We assume this, because for some variables the relation with Uber X waiting time was not significant.

We can confirm H2 for the distribution of the quantity and number of bus lines, which corroborates the results of Lessa et al.'s (2019) study that district area, population density, and percentage of nonwhites are a measure of accessibility. In Hughes and MacKenzie's (2016) and Wang and Mu's (2018) studies, the relationship between Uber's waiting time and the distribution of minorities was not significant.

DISCUSSION AND CONCLUSION

The final model presents different variables from similar studies. Minority rate, for instance, presents a high significance with a p-value of 0.00, in contrast with the cases of Seattle (Hughes & MacKenzie, 2016) and Atlanta (Wang & Mu, 2018). The relevance of local context is important to explain this study's findings. Because waiting times can have a relation to consistent supply and demand of cars (Hall et al., 2019), one possible explanation is that, as the socioeconomic spatial pattern (Torres et al., 2003) also goes along with minority rates and peripheral status, the service tends to be less accessible to these regions in which its demand could be lower. In the case of São Paulo, the correlation between minority rates and income encourage us to at least think on this possibility, as the affordability of Uber can be relatively lower than in the contexts of Seattle or Atlanta. Future analysis using data from Uber cost estimates can help to explore this pattern.

By contrast, some variables such as population density (Hughes & MacKenzie, 2016; Lessa et al., 2019; Wang & Mu, 2018) and road density (Wang & Mu, 2018), find some resonance in the discussion. Even their being far from enough to explain accessibility and being aware of the local nature of the problem (Schwanen, 2017), it is possible to propose them as variables that can help to describe local variables. This possibility exists because their relation to accessibility appears to be consistent.

The need to promote intuitive and highly communicable accessibility measures is undeveloped among researchers n researchers (Páez et al., 2012). Similarly, the construction of a transport accessibility measure from Big Data tools such as Uber's API can contribute to the communication and understanding of this indicator for the general public as the theory gets close to a service of mass consumption. The expansion of the same analysis to other service providers and a more precise identification of this market share in the transportation field in São Paulo can also help to identify possible biases in using Uber's tool.

Public management can make use of similar tools to develop a "basket" of indicators for different modes of transportation and measurements regarding their interrelations. These can be used to better regulate existing activity and to

 CC BY ISSN 0034-7590; eISSN 2178-938X

tailor public service delivery in mobility, as with a more intimate integration between public and private modes. By maintaining a real time measurement of the effect of public regulation in aspects of the transport system, it could be possible to better adjust intervention in the interests of the users (Jin et al., 2018), given that well balanced indicators can reflect user behavior and needs (Lessa et al., 2019; Páez et al., 2012).

However, some limitations of the study deserve attention, as they provide indications for future research. It is important to underline the importance of more comprehensive analysis. This study limited itself to Uber's fleet. In cities like São Paulo where there are at least one more player with a significant fleet, it will be interesting to replicate the methodology with other companies that provide similar services. Second, the replication of the study in other Brazilian cities appears necessary because the economic and cultural factors that affect the disparities may then be analyzed in contexts that are more similar than Atlanta and Seattle, for instance. Third, there is a need to deepen the understanding of Uber's time estimate tool to identify possible biases that could have suppressed any relation. Comparing it to other services can be useful in this sense. A better comprehension of this data can be a major advance in tools for municipalities to better serve their population's transportation needs. Furthermore, it is worth revisiting this study after the publication of the 2020 IBGE Census, because many of the variables used in the study are from the 2010 Census and are projections from it.

REFERENCES

Aranha, V. (2005). *Mobilidade pendular na metrópole paulista*. *São Paulo em Perspectiva*, *19*(4), 96-109. doi:10.1590/S0102-88392005000400006

Batterham, A. M., Tolfrey, K., George, K. P. (1997). Nevill's explanation of Kleiber's 0.75 mass exponent: An artifact of collinearity problems in least squares models? *Journal of Applied Physiology*, *82*(2), 693-697. doi:10.1152/jappl.1997.82.2.693

Cohen, P., Metcalfe, R., Angrist, J., Chen, K., Doyle, J., Farber, H., Hahn, R. (2016). Using big data to estimate consumer surplus: The case of Uber. *National Bureau of Economic Research* (Working paper nº 22627). Retrieved from https://www.nber.org/papers/w22627 doi: 10.3386/w22627

Flores, O., & Rayle, L. (2017). How cities use regulation for innovation: The case of Uber, Lyft and Sidecar in San Francisco. *Transportation Research Procedia*, *25*, 3756-3768. doi:10.1016/j.trpro.2017.05.232

Francisco, E. de R. (2010). *Indicadores de renda baseados em consumo de energia elétrica: abordagens domiciliar e regional na perspectiva da estatística espacial* (Doctoral Thesis, Fundação Getulio Vargas). Retrieved from https://bibliotecadigital.fgv.br/dspace/handle/10438/8158

Hair, J., Babin, B., Money, A., & Samouel, P. (2005). *Fundamentos de métodos de pesquisa em administração*. Porto Alegre, RS: Bookman Companhia Ed.

Hall, J. V, Horton, J. J., & Knoepfle, D. T. (2019). *Pricing efficiently in designed markets: The Case of Ride-Sharing*. 1-76. Retrieved from http://john-joseph-horton.com/papers/uber_price.pdf

Hall, J. V, & Krueger, A. B. (2016). An analysis of the labor market for Uber's driver-partners in the United States. In *National Bureau of Economic Research* (Working paper nº 22843). Retrieved from https://www.nber.org/papers/w22843. doi:10.3386/w22843.

Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Institute of planners, 25*(2), 73-76. doi:10.1080/01944365908978307

Hughes, R., & MacKenzie, D. (2016). Transportation network company wait times in Greater Seattle, and relationship to socioeconomic indicators. *Journal of Transport Geography*, *56*, 36-44. doi:10.1016/j.jtrangeo.2016.08.014

Instituto Brasileiro de Geografia e Estatística. (2010). Censo demográfico. *Notas Metodológicas–Microdados da Amostra*. *Rio de Janeiro*. Retrieved from ftp://ftp.ibge.gov.br/Censos/Censo_Demografico_2010/Resultados_Gerais_da_Amostra/Microdados/Documentacao.zip

Jin, S. T., Kong, H., Wu, R., & Sui, D. Z. (2018). Ridesourcing , the sharing economy , and the future of cities. *Cities*, *76*(January), 96-104. doi:10.1016/j.cities.2018.01.012

Kim, B. G., Trimi, S., & Chung, J. (2014). Big-Data applications in the government sector. *Communications of the ACM*, *57*(3), 78-85. doi:10.1145/2500873

Krajzewicz, D., Erdmann, J., Behrisch, M., & Bieker, L. (2012). Recent development and applications of SUMO-Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements, 5*(3-4), 128-138.

Kwan, M. (1998). Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis, 30*(3), 191-216. doi:10.1111/j.1538-4632.1998.tb00396.x

Kwan, M. P. (2016). Algorithmic geographies: Big data, algorithmic uncertainty, and the production of geographic knowledge. *Annals of the American Association of Geographers, 106*(2), 274-282. doi:10.1080/00045608.2015.1117937

Lessa, D. A., Lobo, C., & Cardoso, L. (2019). Accessibility and urban mobility by bus in Belo Horizonte / Minas Gerais – Brazil. *Journal of Transport Geography*, *77*(May 2018), 1-10. doi:10.1016/j.jtrangeo.2019.04.004

Letouzé, E., & Jütting, J. (2015). Official statistics, Big Data and human development: towards a new conceptual and operational approach. *Data Pop Alliance White Paper Series*. Retrieved from https://paris21.org/sites/default/files/WPS_OfficialStatistics_June2015.pdf

Litman, T. (2003). Measuring transportation: Traffic mobility and accessibility. *Victoria Transport Policy Institute. ITE Journal*, *73*(10), 28-32.

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*. Retrieved from https://hbr.org/

Metrô. (2008). Pesquisa Origem-Destino 2007. *Secretaria de Transportes Metropolitanos*. Retrieved from http://www.metro.sp.gov.br/pesquisa-od/arquivos/OD_2007_Sumario_de_Dados.pdf

Ministério do Trabalho (2016). Relação Anual de Informações Sociais - RAIS ano-base 2016. *Portaria n. 1464 de 30 de Dez. 2016. Aprova Instruções Para a Declaração Da Relação Anual de Informações Sociais - RAIS Ano-Base 2016*. Retrieved from http://www.rais.gov.br/sitio/sobre.jsf

Morandi, E., Ribeiro, R., Hernandes, E., Camara, B., Spinola, L., & Francisco, E. D. R. (2016). *Análise Geoespacial da Relação entre Transporte Público sobre Trilhos, Renda e Tempo Médio de Deslocamento*. Paper presented on XL EnANPAD, Salvador, BA. Retrieved from http://www.anpad.org.br/~anpad/eventos.php?cod_evento=1&cod_edicao_subsecao=1302&cod_evento_edicao=83&cod_edicao_trabalho=21023

Noulas, A., Scellato, S., Lambiotte, R., Pontil, M., & Mascolo, C. (2012). A Tale of many cities: Universal patterns in human urban mobility. *PloS One*, *7*(9). doi:10.1371/annotation/ca85bf7a-7922-47d5-8bfb-bcdf25af8c72

Páez, A., Scott, D. M., & Morency, C. (2012). Measuring accessibility: Positive and normative implementations of various accessibility indicators. *Journal of Transport Geography*, *25*, 141–153. https://doi.org/10.1016/j.jtrangeo.2012.03.016

Pääkkönen, P., & Pakkala, D. (2015). Reference architecture and classification of technologies, products and services for big data systems. *Big data research*, *2*(4), 166-186. doi:10.1016/j.bdr.2015.01.001

Preston, V., & McLafferty, S. (1999). Spatial mismatch research in the 1990s: Progress and potential. *Papers in Regional Science*, *78*(4), 387-402. doi:10.1111/j.1435-5597.1999.tb00752.x

Schwanen, T. (2016). Geographies of transport I: Reinventing a field? *Progress in Human Geography*, *40*(1), 126-137. doi:10.1177/0309132514565725

Schwanen, T. (2017). Geographies of transport II: Reconciling the general and the particular. *Progress in Human Geography*, *41*(3), 355-364. doi:10.1177/0309132516628259

Stelder, D. (2016). Regional accessibility trends in Europe: Road Infrastructure, 1957-2012. *Regional Studies*, *50*(6), 983-995. doi:10.1080/00343404.2014.952721

Torres, H. da G., Marques, E., Ferreira, M. P., & Bitar, S. (2003). Pobreza e espaço: Padrões de segregação em São Paulo. *Estudos Avançados*, *17*(47), 97-128. doi:10.1590/S0103-40142003000100006

Tribby, C. P., & Zandbergen, P. A. (2012). High-resolution spatio-temporal modeling of public transit accessibility. *Applied Geography Journal*, *34*, 345-–355. doi:10.1016/j.apgeog.2011.12.008

Wang, M., & Mu, L. (2018). Spatial disparities of Uber accessibility: An exploratory analysis in Atlanta, USA. *Computers, Environment and Urban Systems*, *67*, 169-175. doi:10.1016/j.compenvurbsys.2017.09.003

Weibull, J. W. (1980). On the numerical measurement of accessibility. *Environment and Planning A: Economy and Space, 12*(1), 53-67. doi:10.1068/a120053

Zhou, X., Wang, M., & Li, D. (2017). From stay to play–A travel planning tool based on crowdsourcing user-generated contents. *Applied Geography, 78*, 1-11. doi:10.1016/j.apgeog.2016.10.002

FORUM

Submitted 07.24.2018. Approved 07.19.2019
Evaluated through a double-blind review process. Guest Scientific Editors: Eduardo de Rezende Francisco, José Luiz Kugler, Soong Moon Kang, Ricardo Silva, and Peter Alexander Whigham
Original version
DOI: http://dx.doi.org/10.1590/S0034-759020190607

FACTORS AFFECTING THE ADOPTION OF BIG DATA ANALYTICS IN COMPANIES

Fatores que afetam a adoção de análises de Big Data em empresas

Factores que afectan a la adopción del análisis Big Data en empresas

ABSTRACT

With the total quantity of data doubling every two years, the low price of computing and data storage, make Big Data analytics (BDA) adoption desirable for companies, as a tool to get competitive advantage. Given the availability of free software, why have some companies failed to adopt these techniques? To answer this question, we extend the unified theory of technology adoption and use of technology model (UTAUT) adapted for the BDA context, adding two variables: resistance to use and perceived risk. We used the level of implementation of these techniques to divide companies into users and non-users of BDA. The structural models were evaluated by partial least squares (PLS). The results show the importance of good infrastructure exceeds the difficulties companies face in implementing it. While companies planning to use Big Data expect strong results, current users are more skeptical about its performance.

KEYWORDS | Big Data, intention behavior, unified theory of acceptance and use of technology, resistance to use, perceived risk.

RESUMO

Com a quantidade total de dados duplicando a cada dois anos, o baixo preço da computação e do armazenamento de dados tornam a adoção de análises de Big Data (BDA) desejável para as empresas, como aquelas que obterão uma vantagem competitiva. Dada a disponibilidade de software livre, por que algumas empresas não adotaram essas técnicas? Para responder a essa pergunta, estendemos a teoria unificada de adoção e uso de tecnologia (UTAUT) adaptado para o contexto do BDA, adicionando duas variáveis: resistência ao uso e risco percebido. Usamos a nível da implementação da tecnologia para dividir as empresas em usuários e não usuários de técnicas de BDA. Os modelos estruturais foram avaliados por partial least squares (PLS). Os resultados mostram que a importância de uma boa infraestrutura excede as dificuldades que as empresas enfrentam para implementá-la. Enquanto as empresas que planejam usar Big Data esperam resultados fortes, os usuários atuais são mais céticos em relação ao seu desempenho.

PALAVRAS-CHAVE | *Big Data, intenção de usar, teoria unificada de adoção e uso de tecnologia, resistência ao uso, risco percebido.*

RESUMEN

Con la cantidad total de datos duplicándose cada dos años, el bajo precio de la informática y del almacenamiento de datos, la adopción del análisis Big Data (BDA) es altamente deseable para las empresas, como un instrumento para conseguir una ventaja competitiva. Dada la disponibilidad de software libre, ¿por qué algunas empresas no han adoptado estas técnicas? Para responder a esta pregunta, ampliamos la teoría unificada de la adopción y uso de tecnología (UTAUT) adaptado para el contexto BDA, agregando dos variables: resistencia al uso y riesgo percibido. Utilizamos el grado de implantación de estas técnicas para dividir las empresas entre: usuarias y no usuarias de BDA. Los modelos estructurales fueron evaluados con partial least squres (PLS). Los resultados muestran que la importancia de una buena infraestructura excede las dificultades que enfrentan las empresas para implementarla. Mientras que las compañías que planean usar BDA esperan muy buenos resultados, las usuarias actuales son más escépticos sobre su rendimiento.

PALABRAS CLAVE | *Big Data, intención de uso, teoría unificada de la adopción e uso de tecnología, resistencia al uso, riesgo percibido.*

JUAN-PEDRO CABRERA-SÁNCHEZ[1]
jcabrera10@us.es
ORCID: 0000-0001-5723-3153

ÁNGEL F. VILLAREJO-RAMOS[1]
curro@us.es
ORCID: 0000-0002-6916-2839

[1]Universidad de Sevilla, Facultad de Ciencias Económicas y Empresariales, Sevilla, Spain

INTRODUCTION

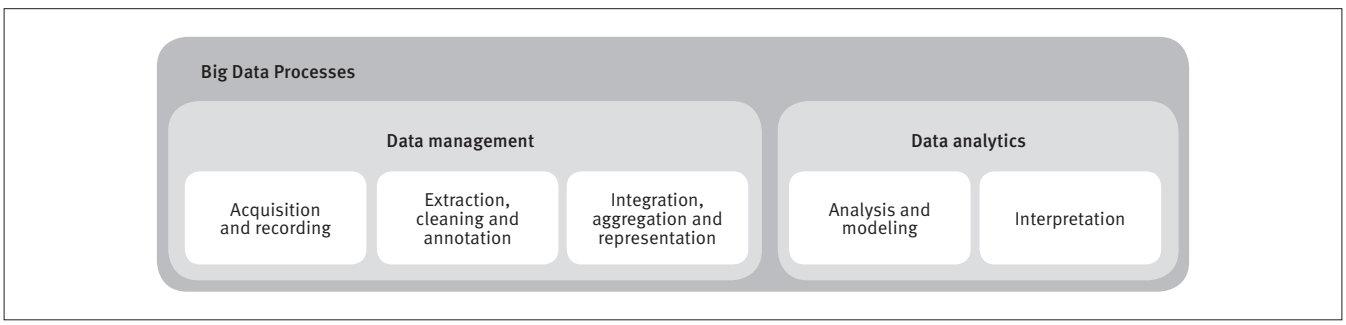
Currently, society generates data about our activities at an exponential rate of growth. This data covers, for example, our mobile phones and their location, any online transactions, the Internet of things, social networks, wearables, etc. Firms that can transform these data into real-time information about their customers gain a substantial competitive advantage (Sivarajah, Kamal, Irani & Weerakkody, 2016). User data allows firms to know when their customers consume their products, the best times for promotions, and how to improve brand sentiments. Firms using Big Data analytics (BDA) (McAfee & Brynjolfsson, 2012) can process huge quantities of data, almost in real time, and become leaders in the market.

The adoption, implementation, and management of BDA requires companies to acquire new skills. New career profiles such as data scientist, which combines engineering, statistics, and a deep knowledge of business, are among the most sought-after jobs nowadays. Employees with these skills help companies mine data generated by the companies themselves and their customers. This changes how decisions are made, favoring a data-driven approach over one based on the personal experience of CEOs (McAfee & Brynjolfsson, 2012).

Firms using BDA must address the challenges that arise in the so-called data cycle of life: questions about the data themselves, difficulties processing the data, and concerns about data management (Akerkar, 2014; Zicari, 2014). Questions about the data themselves revolve around their volume, variety, velocity, veracity, volatility, value, and visualization. Data processing tasks include the techniques related to data acquisition, storage in databases, cleaning and transforming existing data, correct model selection, and presentation of the results. Finally, proper data management involves ethical considerations, including respect for user privacy and security.

Because decision-making is increasingly data-driven, companies must obtain valuable information in an efficient way from a rapidly changing data environment. This process, detailed by Agrawal, Bernstein, & Bertino (2011) as it shows in Figure 1.

Figure 1. **Big Data processes**



As Figure 1 shows, companies' use of BDA involves two major processes: data management and data analytics. While data management raises questions of engineering, data analytics speaks more directly to our interests as marketers. BDA is the process of getting value from data by finding hidden patterns that support data-driven decision making.

Companies considering BDA adoption face several barriers such as lack of knowledge, fear, resistance to change, and the technology's own limitations (Yaqoob et al., 2016). However, BDA improves their decision-making, utilizing techniques and software that are free and open source. This leads us to ask two questions. First, what affects its adoption? Second, why are there many companies that do not use BDA yet? Most of the literature on BDA focuses on technical aspects related to its ecosystem: application development, data mining, analytics, prediction, prescription, or statistical modeling (Sivarajah et al., 2016). There is little research about BDA adoption inside companies (Kwon, Lee, & Shin, 2014; Brünink, 2016; Rahman, 2016; Demoulin & Coussement, 2018; Huang, Liu, & Chang, 2012; Verma, Bhattacharyya, & Kumar, 2018).

This study, based on the unified theory of acceptance and use of technology (UTAUT) model (Venkatesh, Morris, Davis, & Davis, 2003), considers the impact of two new variables, resistance to use and perceived risk, on the adoption of BDA. The aim of this study is to explain the adoption and use of this new technology by companies and to understand implementation problems in order to give recommendations to practitioners. This is why we differentiate between user and non-user companies of this technology and look for different factors that affect its acceptance and use.

The second section of the paper describes the theoretical foundations of the proposed model. The third section describes the methodology we use. The fourth section analyzes the results

obtained from the application of the model to a sample of companies. Finally, we present the main theoretical and practical conclusions, as well as the limitations, of our research.

THEORETICAL BACKGROUND AND HYPOTHESES OF THE PROPOSED MODEL

Big Data analytics and marketing

Big Data has revolutionized marketing analytics and marketing in general. It has provided new concepts and new ways of doing things (Watson, 2019; Wedel & Kannan, 2016) to generate a competitive advantage. BDA enables service innovation that creates strategic value for companies (Chiang, Grover, Liang, & Zhang, 2018).

Currently, BDA is being used for marketing campaigns oriented toward increasing customer engagement (Liu, Shin, & Burns, 2019). The literature review shows an increasing number of publications on the use of Big Data techniques aimed at creating relational marketing benefits. (Amado, Cortez, Rita, & Moro, 2018).

Marketing management can use the huge amount of data available (for example, in social media) to gain valuable insights from their customers. Companies that exploit Big Data from social media gain competitive advantages because they know customers better (Ducange, Pecori, & Mezzina, 2018). Studies show that, using BDA for business intelligence (Sun, Sun, & Strang, 2018) and to maintain customer privacy (Palmatier & Martin, 2019) creates important assets in relationship marketing.

However, the literature of BDA adoption is relatively sparse, and it is focused on the industry level (Rehman, Chang, Batool, & Wah, 2016; Wright, Robin, Stone, & Aravopoulou, 2019; Yadegaridehkordi et al., 2018; Lai, Sun, & Ren, 2017). Few authors have researched which factors affect BDA adoption in companies.

Acceptance models of Big Data analytics

Technology adoption by companies and consumers is critical for success. Numerous models of technology adoption have been developed and tested, including the theory of planned behavior (TPB) (Ajzen, 1991) and the technology adoption model (TAM) (Davis, 1985). But, without doubt, the UTAUT model (Venkatesh et al., 2003) is the most comprehensive model. This model integrates previous models and theories in order to analyze technology adoption and acceptance.

Previous studies of BDA adoption in companies (Kwon et al., 2014; Brünink, 2016; Rahman, 2016; Demoulin & Coussement, 2018; Huang et al., 2012; Verma et al., 2018) have used the original TAM (Davis, 1985), TAM2 (Venkatesh & Davis, 2000), TAM3 (Venkatesh & Bala, 2008), or the UTAUT model without any added variables. Acceptance models have been upgraded since their introduction and have even evolved into new models. Because the UTAUT model is already a mature model, we enhance it with two new variables (discovered to be significant in this research), which help explain whether companies choose to adopt BDA.

The behavioral reasoning theory (Claudy, Garcia, & O'Driscoll, 2015) provides a framework in which user involvement is very important for successful technology adoption (Ives & Olson, 2008). Users who are predisposed to change have less resistance to adopting a new technology (Laumer, Maier, Eckhardt, & Weitzel, 2016). Different attitudes shape the adoption process of a new technology (Gargallo López, Suárez Rodríguez, & Almerich Cerveró, 2006). This research led us to search for different patterns among users or non-users in our sample of companies.

Our proposed model includes four independent variables drawn from the UTAUT model. First, performance expectancy is defined as the degree to which the use of technology is expected to offer benefits for the company. Second, effort expectancy measures the ease of use expected for a technology. Third, social influence measures how individuals perceive that friends and family think that they should use a technology. Fourth, facilitating conditions is defined as the extent to which consumers perceive that resources and support will be available to develop a behavior. The model proposes a direct influence of the first three variables on behavioral intention, while facilitating conditions affects behavioral intention and usage behavior. Arenas-Gaitán, Peral-Peral, and Villarejo-Ramos (2016) indicate that the value of this model is its capacity to identify which factors are the main determinants of adoption. The model allows the inclusion of different moderating variables that affect the influence of the model's key constructs.

We add resistance to use and perceived risk to the UTAUT constructs. Resistance to use consists of negative reactions to change or new system implementation (Kim & Kankanhalli, 2009). Perceived risk is the potential for losses as a result of the implementation of a new technology or information system (Featherman & Pavlou, 2003).

Hypotheses of the proposed model

We propose several hypotheses based on the extended UTAUT model for the acceptance and use of BDA in companies.

 CC BY ISSN 0034-7590; eISSN 2178-938X

Performance expectancy refers to the perception of the performance that the technology is going to have and is one of the most influential constructs regarding behavioral intention. Several studies (Brünink, 2016; Chauhan & Jaiswal, 2016; Yu, 2012) besides the original work (Venkatesh et al., 2003) support this positive relationship. Therefore, we propose as a hypothesis:

> H1: Performance Expectancy positively influences the behavioral intention to use BDA.

Effort expectancy refers to the ease of learning and use of this new technology. According to the UTAUT model, the degree to which BDA will be adopted depends on the ease or difficulty of its use. Several studies find support for this relationship (Al-Gahtani, Hubona, & Wang, 2007; Chauhan & Jaiswal, 2016; Kim, Chan, & Gupta, 2007; Lee & Song, 2013; Yu, 2012) and confirm the effect of effort expectancy on behavioral intention. Thus, we offer as a second hypothesis of the model:

> H2: Effort Expectancy, or ease of use, positively affects the behavioral intention to use BDA.

The concept of social influence developed in the original work of Venkatesh et al. (2003) and extended in UTAUT2 (Venkatesh, Thong, & Xu, 2012) measures the effect of what others (friends and family) think about this technology. In a professional environment, what managers and colleagues think is also very important (Al-Gahtani et al., 2007; Brünink, 2016; Chauhan & Jaiswal, 2016; Gupta, Huang, & Niranjan, 2010; H. W. Kim et al., 2007; Lee & Song, 2013). Therefore, we propose as a hypothesis:

> H3: Social Influence positively affects the behavioral intention to Use BDA.

Resistance to use consists of opposition or negative reactions to the implementation of a new technology. As Gibson (2004) finds, the introduction of many new technologies have failed due to the opposition of users to their implementation. Although current literature recognizes resistance to use (Kim & Kankanhalli, 2009; Lapointe & Rivard, 2007), there are few studies that integrate it into the UTAUT model. Nevertheless, there are precedents for using it to explain behavioral intention (Hsieh, 2015). Norzaidi, Salwani, Chong, and Rafidah (2008) verify the relationship between user resistance and usage, a finding confirmed by other studies that do not use the UTAUT model (Bhattacherjee & Hikmet, 2007; Poon et al., 2000). Therefore, we offer as a hypothesis:

> H4: Resistance Use negatively affects behavioral intention to use BDA.

Perceived risk consists of the potential for losses in the implementation of a new technology. In addition to the work of Featherman and Pavlou (2003) which includes the measurement scale that we use, many studies about perceived risk as a negative antecedent of behavioral intention (Kim, Ferrin, & Rao, 2008; Lee & Song, 2013; Martins, Oliveira, & Popovič, 2014). Therefore, we propose as a hypothesis:

> H5: Perceived Risk negatively affects behavioral intention to use BDA.

Facilitating conditions are favorable when there is easy access to the resources needed to use a new technology and to subsequent support (Venkatesh et al., 2003). In later studies using UTAUT2, Venkatesh et al. (2012) found that this construct has a significant effect on the behavioral intention to use a new technology. Also, more recent studies have verified this positive effect on behavioral intention (Duyck et al., 2010; Hung, Wang, Cho, & Chou, 2007; Wu, Tao, & Yang, 2007). Thus, we offer as a hypothesis:

> H6: Facilitating Conditions positively influence the behavioral intention to use BDA.

Both TPB (Ajzen, 1991) and UTAUT (Venkatesh et al., 2003) have been used to show how favorable facilitating conditions positively affect the use of a new technology. Various subsequent works (Al-Gahtani et al., 2007; Brünink, 2016; Chauhan & Jaiswal, 2016; Duyck et al., 2010; Kim et al., 2007) also support this relationship. Therefore, we propose as a hypothesis:

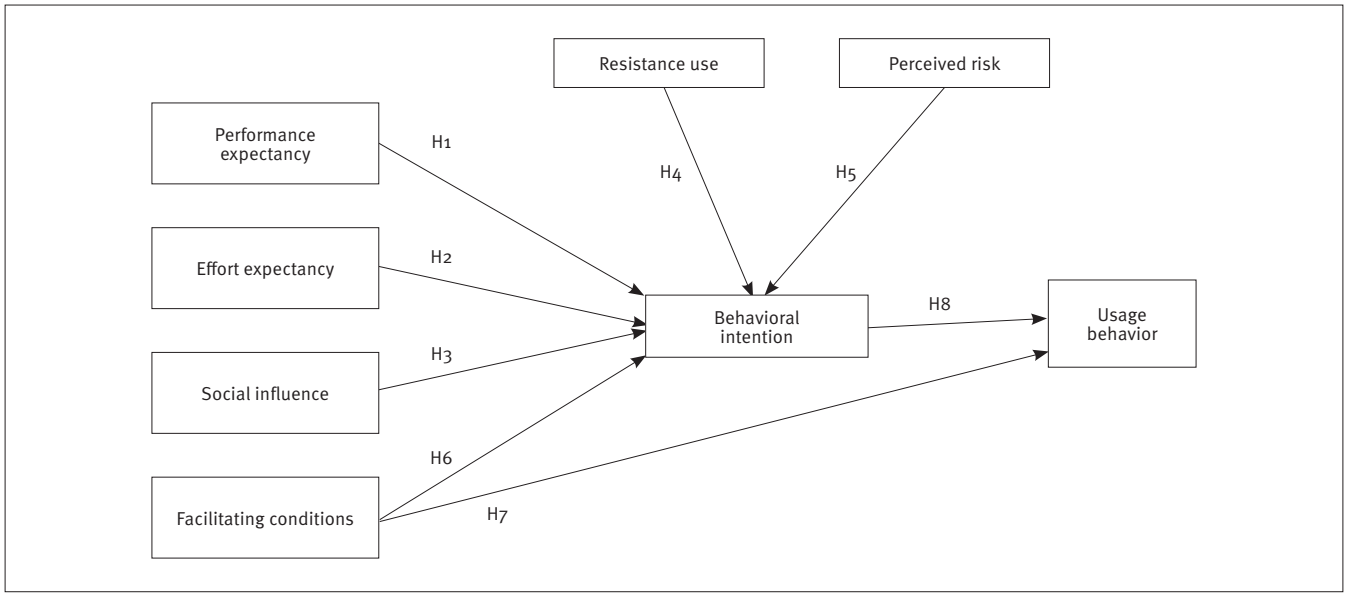
> H7: Facilitating Conditions positively affect the use of BDA.

The main technology acceptance models (TRA, TAM, UTAUT, and UTAUT2) show a direct relationship between behavioral intention and the use of technologies (Fishbain & Ajzen, 1975; Davis, 1985; Venkatesh et al., 2003; 2012). This influence has been demonstrated in contexts similar to the adoption of BDA, such as internet banking (Martins et al., 2014), online flight purchasing (Escobar-Rodríguez & Carvajal-Trujillo, 2014), electronic document management systems (Afonso, Gonzalez, Roldán, & Sánchez-Franco, 2012) and ERP (Enterprise Resource Planning) (Chauhan & Jaiswal, 2016). Therefore, we can enunciate as a hypothesis:

> H8: The behavioral intention to use BDA positively affects its use.

In Figure 2, we show the proposed model of the acceptance and use of BDA with pathways identified in our hypotheses.

Figure 2. **Big Data Acceptance model in companies**



RESEARCH METHODOLOGY

Sample description

Our survey, collected between September and October of 2017 via personal emails and phone, sampled 199 responses of company CEOs and managers of different areas, such as human resources, finance, marketing, and sales. A pre-test of the survey was carried out with five volunteer managers and several expert researchers, who completed the questionnaire and provided feedback about the questions. In Table 1, we classify the companies of the respondents according to their revenues and sectors.

Table 1. **Companies of the sample according to revenue and activity sector**

	‹ 2M€	2M€‹X‹10M€	10M€‹X‹43M€	› 43 M€	(not answered)	Total
Agriculture	1	3	2	1		7
Commerce and distribution	5	4	1	10		20
Telco	6	2	4	14	1	27
Construction	2		1	4		7
Education	2	1		2		5
Energy	1			3		4
Finance	1		2	8		11
Industrial	5	3	2	6		16
Others	10	10	6	13	2	41
Health	3			2		5
Services	24	12	9	10		55
(not answered)					1	1
Total general	**60**	**35**	**27**	**73**	**4**	**199**

Measurement scales

For the UTAUT constructs, we adapt scales from Venkatesh et al. (2003) to BDA. Resistance to use was measured with the scale proposed by Bhattacherjee and Hikmet (2007), while perceived risk was measured with the scale from Featherman and Pavlou (2003). Seven point Likert scales were used in all cases.

Statistical tools

To estimate the structural model, we used partial least squares (PLS), (Chin & Dibbern, 2010; Hair et al., 2012) with the statistical software Smart PLS 3.2.3 (Ringle, Wende, & Becker, 2015). To avoid measurement bias, or common method bias (CMB), in the observed sample, we followed the recommendations of Burton-Jones (2009). We have also follow Podsakoff, MacKenzie, Lee, & Podsakoff (2003), MacKenzie, Podsakoff, & Podsakoff (2011); Podsakoff, MacKenzie, & Podsakoff (2012), and Kock & Lynn (2012). Because the study focuses on CMB and PLS for structural equation models. We follow Kock (2015) and add unrelated questions in order to create a new latent variable with these indicators and the other variables as antecedents. This CMB variable acts as the dependent variable for all of the others in the model. The variance inflation factors estimated by this method must be lower than 3.3 to confirm that the sample does not have CMB. In Table 2, we show that our sample complies with this requirement.

Table 2. **VIF from all variables to check CMB**

	Variable_CMB
Behavioral intention	2.423
Effort expectancy	1.631
Facilitating conditions	2.472
Perceived risk	1.288
Performance expectancy	1.994
Resistance to use	1.852
Social influence	1.675
Usage behavior	1.996

RESULTS

We checked the reliability of all of the constructs. Current literature suggests that for measurement models to be considered reliable and valid, each factor loading should exceed 0.7 (Roldán & Sánchez-Franco, 2012; Henseler, Ringle, & Sarstedt, 2014). In Table 3, we show that each loadings was over 0.7, except for the third indicator for facilitating conditions (FC3), which was dropped.

Table 3. **Reliability of measurement scales (loadings)**

	Behavioral intention	Effort expectancy	Facilitating conditions	Perceived risk	Performance expectancy	Resistance to use	Social influence	Usage behavior
BI1	0.970							
BI2	0.986							
BI3	0.984							
BI4	0.974							
EE1		0.777						
EE2		0.887						
EE3		0.898						
EE4		0.874						
EE5		0.806						
FC1			0.874					
FC2			0.892					
FC4			0.847					
PR1				0.927				

Continue

 CC BY ISSN 0034-7590; eISSN 2178-938X

Table 3. **Reliability of measurement scales (loadings)** Conclusion

	Behavioral intention	Effort expectancy	Facilitating conditions	Perceived risk	Performance expectancy	Resistance to use	Social influence	Usage behavior
PR2				0.924				
PR3				0.877				
PE1					0.899			
PE2					0.868			
PE3					0.909			
PE4					0.918			
PE5					0.785			
PE6					0.829			
PE7					0.784			
RU1						0.931		
RU2						0.964		
RU3						0.938		
RU4						0.902		
SI1							0.741	
SI2							0.877	
SI3							0.817	
SI4							0.792	
SI5							0.724	
UB								1.000

Next, we analyzed construct reliability using composite reliability indicators and Cronbach's Alpha. In all cases, the values of our indicators were above 0.7 as suggested by Nunnally (1978). We assured convergent validity by analyzing the average variance extracted. All of the values were above the 0.5 threshold proposed by Straub, Detmar, Boudreau, and Gefen (2004). These indicators, shown in Table 4, meet the requirements.

Table 4. **Composite reliability and Convergent validity**

	Cronbach's Alpha	rho_A	Composite reliability	Average Variance Extracted (AVE)
Behavioral intention	0.985	0.986	0.989	0.958
Effort expectancy	0.906	0.934	0.928	0.722
Facilitating conditions	0.841	0.843	0.904	0.759
Perceived risk	0.896	0.909	0.935	0.828
Performance expectancy	0.940	0.947	0.951	0.736
Resistance to use	0.951	0.954	0.965	0.872
Social influence	0.851	0.874	0.893	0.627
Usage behavior	1.000	1.000	1.000	1.000

 CC BY ISSN 0034-7590; eISSN 2178-938X

Next, we evaluated the discriminant validity of the measurement model in two ways. First, we conducted Fornell and Larcker's test (Barclay, Higgins, & Thompson, 1995). Second, using the Heterotrait-Monotrait (HTMT) ratio (Henseler et al., 2014), we ensured that in all cases all of the values were below 0.9. The results of both tests are shown in Tables 5 and 6.

Table 5. **Discriminant validity (Fornell-Larcker's test)**

	Behavioral intention	Effort expectancy	Facilitating conditions	Perceived risk	Performance expectancy	Resistance to use	Social influence	Usage behavior
Behavioral intention	**0.979**							
Effort expectancy	0.384	**0.850**						
Facilitating conditions	0.628	0.587	**0.871**					
Perceived risk	-0.331	-0.185	-0.283	**0.910**				
Performance expectancy	0.544	0.434	0.373	-0.195	**0.858**			
Resistance to use	-0.506	-0.258	-0.343	0.408	-0.566	**0.934**		
Social influence	0.497	0.459	0.483	-0.246	0.479	-0.234	**0.792**	
Usage behavior	0.630	0.361	0.624	-0.276	0.402	-0.400	0.449	**1.000**

Table 6. **Discriminant validity (Ratio Heterotrait-Monotrait -HTMT)**

	Behavioral intention	Effort expectancy	Facilitating conditions	Perceived risk	Performance expectancy	Resistance to use	Social influence	Usage behavior
Behavioral intention								
Effort expectancy	0.380							
Facilitating conditions	0.690	0.649						
Perceived risk	0.349	0.202	0.324					
Performance expectancy	0.559	0.449	0.411	0.206				
Resistance to use	0.521	0.269	0.383	0.443	0.597			
Social influence	0.532	0.507	0.567	0.297	0.524	0.254		
Usage behavior	0.635	0.355	0.679	0.287	0.405	0.408	0.476	

The values of R^2 for the second order constructs (behavioral intention and usage behavior) are shown in Table 7.

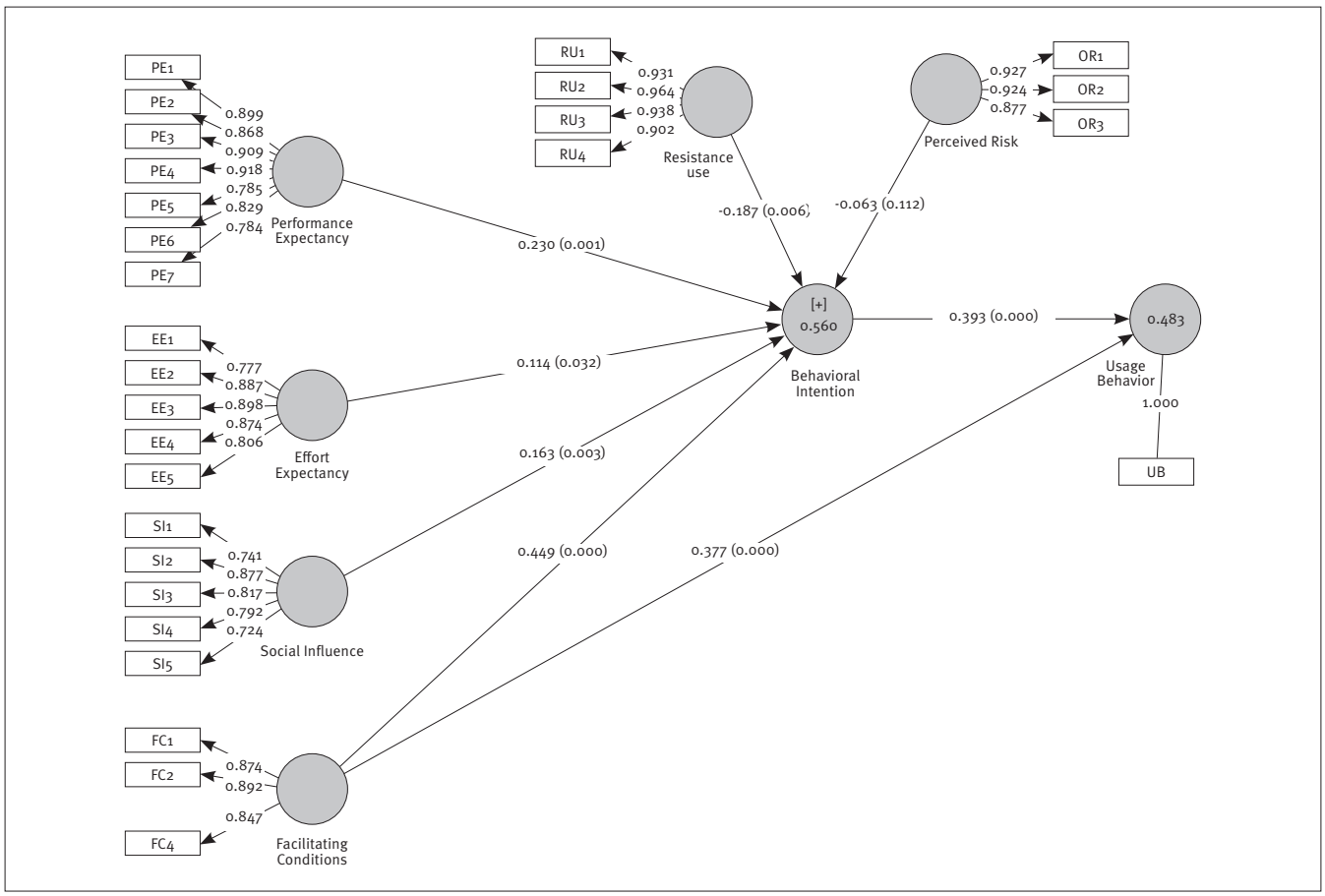
Table 7. **R^2 of the model**

	R squared	Adjusted R squared
Behavioral intention	0.560	0.546
Usage behavior	0.483	0.478

 CC BY ISSN 0034-7590; eISSN 2178-938X

Figure 3 shows the values for every loading and path of the model.

Figure 3. **Results of the model**



To evaluate the structural model, we analyzed the values of the path coefficients and the explained variance of the endogenous variables (R^2). The path coefficients show the intensity of the relationship between the independent and dependent variables. We used a bootstrapping technique with 5,000 samples to find the reliability of the estimated path coefficients, as shown in Table 8.

Table 8. **Structural model estimates (Path coefficients)**

Whole sample	Original sample (O)	P values
Behavioral intention -› Usage behavior	**0.393 *****	**0.000**
Effort expectancy -› Behavioral intention	**-0.114 ***	**0.032**
Facilitating conditions -› Behavioral intention	**0.449 *****	**0.000**
Facilitating conditions -› Usage behavior	**0.377 *****	**0.000**
Perceived Risk -› behavioral intention	-0.063 (n .s.)	0.112
Performance expectancy -› Behavioral intention	**0.230 *****	**0.001**
Resistance to use -› Behavioral intention	**-0.187 ****	**0.006**
Social influence -› Behavioral intention	**0.163 ****	**0.003**

***p‹0.001, **p‹0.01, *p‹0.05. (bootstrapping with 5,000 sub-samples and 1-tailed test).
Significant relationships with path coefficients and p value in bold.

CC BY

We calculated the SRMR (Standardized Root Mean-square Residual) indicator to assess model fit. The value obtained, 0.065, was less than the 0.08 threshold proposed by Henseler et al. (2014), which suggests a good fit for the model. The model explains 47.85% of the variation in usage and 54.6% of the variation in behavioral intention (see Table 7), both of which exceed the minimum level of 10% recommended by Falk and Miller (1992).

The results support most of the hypotheses, except for H5 (perceived risk) and H2 (effort expectancy). The coefficients for supported hypothesis are significant at the 1% level. Although effort expectancy is significant at the 5% level, we have found a negative relationship that is due to a suppressor effect (Falk & Miller, 1992) produced by the new relationship facilitating conditions on behavioral intention so we can reject H2. In order of influence, we can see that facilitating conditions contributes most to behavioral intention and the second most to usage. The second contributor to behavioral intention is performance expectancy while the first contributor to usage is behavioral intention. We also note that the coefficient for behavioral intention on usage is significant at the 0.1% level. The coefficient for the effect of resistance to use on behavioral intention is significant and negative.

We also calculated the Stone-Geisser Q^2 to evaluate the predictive capacity of the model (Gefen, Rigdon, & Straub, 2011). The results are shown in Table 9.

Table 9. **Prediction of latent variables**

	RMSE	Q2
Behavioral intention	0.558	0.502
Usage behavior	0.522	0.397

We conclude that the model has predictive relevance as the values of Q^2 in Table 9 are greater than zero (Roldán & Sánchez-Franco, 2012).

We considered the possibility of heterogeneity in the sample. Following Becker, Rai, Ringle, and Völckner (2013), we ran a PLS-POS latent class segmentation and also a FIMIX latent class segmentation. We find no differences in groups with *a posteriori* segmentation.

Next, we tried several *a priori* segmentations with different criteria (e.g., company size, use of Big Data, activity sector, finding no differences between these sub-samples. However, we did find different behaviors in companies when we asked about the maturity level of the implementation of BDA. We used the scale proposed by Paulk, Curtis, Chrissis and Weber (1993), which has been widely used (Berg, Leinonen, Leivo, & Pihlajamaa, 2002; Khatibian, Hasan, & Jafari, 2010; Urwiler & Frolick, 2008) and has five levels: initial, repeatable, defined, managed, and optimizing. We assigned companies that had not implemented BDA or were in the first two levels to Segment 1 and those in the last three levels to Segment 2. As shown in Table 10 for Segment 1 and in Table 11 for Segment 2, there are significant differences between these two segments and the whole sample (see Table 8).

Table 10. **Segment 1. Structural Model Estimates (Path Coefficients)**

NON-USERS and BEGINERS Segment 1. Size: 152 companies	Original sample (O)	P values
Behavioral intention -› Usage behavior	**0.387 *****	**0.000**
Effort expectancy -› Behavioral intention	-0.092 (n. s.)	0.094
Facilitating conditions -› Behavioral intention	**0.344 *****	**0.000**
Facilitating conditions -› Usage behavior	**0.237 ****	**0.002**
Perceived risk -› Behavioral intention	-0.073 (n. s.)	0.106
Performance expectancy -› Behavioral intention	**0.330 *****	**0.000**
Resistance to use -› Behavioral intention	**-0.227 ****	**0.002**
Social influence -› Behavioral intention	**0.189 ****	**0.002**

***p<0.001, **p<0.01, *p<0.05. (bootstrapping with 5,000 sub-samples and 1-tailed test). Significant relationships with path coefficients and p value in bold.

 CC BY ISSN 0034-7590; eISSN 2178-938X

Table 11. **Segment 2. Structural model estimates (Path coefficients)**

USERS and HEAVY USERS Segment 2. Size: 47 companies	Original sample (O)	P values
Behavioral intention -› Usage behavior	**0.561 *****	**0.001**
Effort expectancy -› Behavioral intention	-0.186 (n. s.)	0.084
Facilitating conditions -› Behavioral intention	**0.648 *****	**0.001**
Facilitating conditions -› Usage behavior	0.128 (n .s.)	0.188
Perceived risk -› Behavioral intention	-0.188 (n. s.)	0.106
Performance expectancy -› Behavioral intention	-0.214 (n. s.)	0.174
Resistance to use -› Behavioral intention	-0.207 (n. s.)	0.172
Social influence -› Behavioral intention	0.056 (n. s.)	0.354

***p<0.001, **p<0.01, *p<0.05. (bootstrapping with 5,000 sub-samples and 1-tailed test). Significant relationships with path coefficients and p value in bold.

We performed an ANOVA test and confirmed the significant differences between the sub-samples. For companies that were non-users or beginners (Segment 1), all relationships were significant except for the effects of perceived risk and effort expectancy on behavioral intention. Facilitating conditions and performance expectancy were the strongest contributors to behavioral intention with high levels of significance. The coefficients for resistance to use and social influence were also large and significant. For users and heavy users (Segment 2), none of the relationships was significant except the effect of facilitating conditions on behavioral intention. It had the strongest effect of all the relationships in this study. Curiously, performance expectancy was not significant.

DISCUSSION, CONCLUSIONS AND LIMITATIONS

Our research extends the UTAUT model for Big Data with a new variable, resistance to use. With this extension, we contribute to the generalization of the model and to a better understanding of technology acceptance. Our model adds to previous research on BDA by including a new independent variable, resistance to use, to the UTAUT model and including usage behavior of BDA as an outcome variable. Brünink (2016) uses the UTAUT model without adding resistance to use or explaining actual usage behavior. Other studies (Demoulin & Coussement, 2018) focus on management support for the use of Big Data applications, using models such as TAM, TAM2, or TAM3 (Brown & Venkatesh, 2005; Huang et al., 2012; Verma et al., 2018) These models explain the adoption and actual use of BDA in companies, but are more limited than the UTAUT model.

Our results show that the behavioral intention to use BDA in companies is determined by four factors. First, performance expectancy, the perception that implementation of this technology will achieve good results, increases adoption, as shown in previous studies (Lee & Song, 2013; Yu, 2012). Second, social influence has a positive effect on the intention to use BDA, as demonstrated in previous papers (Bozan, Parker, & Davey, 2016). Third, facilitating conditions, company provision of support and necessary resources for usage, increases both behavioral intention and usage (Alharbi, 2014). Finally, resistance to use decreases behavioral intentions to use BDA in companies, with a stronger effect than social influence.

We also find that although the use of BDA is perceived to be difficult (effort expectancy), the influence of this perception on behavioral intention is small and contained into other relationship: facilitation conditions on behavioral intention (suppressor effect abovementioned).

We also find a positive effect of facilitating conditions on the usage behavior of the new technology with a similar loading for behavioral intention. Thus, we can say that the findings are consistent with all of the hypotheses of the UTAUT model, except for H5 (perceived risk). Because we find it has a significant effect, we propose adding resistance to use to the original model, in order to improve explanations of the acceptance and use of BDA in companies.

Finally, we highlight the differences in behavior between companies that are not using these techniques or are beginning to use them (Segment 1) and the companies that have already been using them for a long time (Segment 2). For the beginners or non-user companies, performance expectancy, social influence, and facilitating conditions have strong positive effects on behavioral intention and usage behavior, while resistance to use has a strong negative effect on both variables. Among user and heavy user companies, only facilitating conditions has an effect on behavioral intention, while the rest of relationships are not significant. This may suggest that established users know what they can achieve with these techniques so the only thing they care is about having good facilitating conditions while beginners still do not know

 CC BY © RAE | São Paulo | **59(6)** | **November-December 2019** | 415-429 ISSN 0034-7590; eISSN 2178-938X

all the capabilities of this technology so they take into account more issues.

As to professional implications, the results suggest that executives assume that every technology has its own learning curve, and this issue does not affect its adoption whenever great results are expected, as in the case of Big Data (Cabrera-Sánchez & Villarejo-Ramos, 2018). Likewise, if the company has the appropriate infrastructure, it does not lose anything if it tests the technology. In any case, overcoming the resistance to use it still requires clear information about its benefits. Therefore, we recommend two steps for managers. First, they should be informed that most software associated with these techniques is free and that if they already have hardware resources, they should test it. Second, there should be communication with managers about the benefits of using Big Data, including examples of companies using it in the same areas. This second action is very important for companies who are currently using Big Data because we can infer that they are not squeezing all of the potential from the technology. They are somewhat upset about the technology, and they have a very poor performance expectancy when this just the opposite should be true. Therefore, we must inform them about the technology and how it can be used to generate profits in each sector.

The use of BDA in companies can be a very important advance in information management to improve customer relationships. As it is more than a customer relationship management tool, BDA gives companies relevant information and increases customers' knowledge, improving their engagement.

Although the UTAUT model is well-tested and mature, we have included two variables to extend it. However, there may be many more variables that are relevant. For this technology, there are constructs from the original model such as performance expectancy that have a lower influence on behavioral intention than one of the constructs we added, resistance to use. Because other new variables with strong effects may exist, the UTAUT model must continue to evolve in order to provide better explanations for the acceptance of new technologies. Future research on Big Data should seek to identify these variables. Also, it seems necessary to explore new moderator variables with the aim of analyzing possible effects not previously taken into account.

Finally, larger sample sizes will allow us to establish differences in behavior between groups of companies, which we can analyze via an *a posteriori* segmentation technique, such as Posteriori Oriented Segmentation- Partial Least Squares (POS-PLS). So, if we get a bigger sample of companies that are using (or are intending to use) Big Data, we will have a better performance of this model and more informative results.

REFERENCES

Afonso, C., Gonzalez, M., Roldán, J., & Sánchez-Franco, M. (2012). Determinants of user acceptance of a local eGovernment Electronic Document Management System (EDMS). In *Proceedings of the 12th European Conference on e-Government, ECEG* (pp. 19-28), Barcelona: Academic Publishing International Limited.

Agrawal, D., Bernstein, P., & Bertino, E. (2011). Challenges and opportunities with Big Data 2011-1. *Proceedings of the VLDB Endowment* (pp. 1-16). Retrieved from http://dl.acm.org/citation.cfm?id=2367572%5Cnhttp://docs.lib.purdue.edu/cctech/1/

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179-211. doi:10.1016/0749-5978(91)90020-T

Akerkar, R. (2014). Analytics on big aviation data: Turning data into insights. *International Journal of Computer Science and Applications*, *11*(3), 116-127.

Al-Gahtani, S. S., Hubona, G. S., & Wang, J. (2007). Information technology (IT) in Saudi Arabia: Culture and the acceptance and use of IT. *Information & Management*, *44*(8), 681-691. doi:10.1016/j.im.2007.09.002

Alharbi, S. T. (2014). Trust and acceptance of cloud computing: A revised UTAUT model. *Proceedings - 2014 International Conference on Computational Science and Computational Intelligence, CSCI 2014*, *2*(Mm) (pp. 131-134). doi:10.1109/CSCI.2014.107

Amado, A., Cortez, P., Rita, P., & Moro, S. (2018). Research trends on Big Data in marketing: A text mining and topic modeling based literature analysis. *European Research on Management and Business Economics*, *24*(1), 1-7. doi:10.1016/j.iedeen.2017.06.002

Arenas-Gaitán, J., Peral-Peral, B., & Villarejo-Ramos, A.-F. (2016). Grupos de mayores en la banca electrónica. Segmentación de clases latentes con PLS-POS. In *Congreso Marketing AEMARK*. Madrid, Spain.

Barclay, D., Higgins, C., & Thompson, R. (1995). The Partial Least Squares (PLS) approach to causal modelling: Personal computer adoption and use as an illustration. *Technology Studies,* 2(2), 285-309.

Becker, J.-M., Rai, A., Ringle, C. M., & Völckner, F. (2013). Discovering unobserved heterogeneity in structural equation models to avert Validity threats. *MIS Quarterly*, *37*(3), 665-694. doi:10.25300/MISQ/2013/37.3.01

Berg, P., Leinonen, M., Leivo, V., & Pihlajamaa, J. (2002). Assessment of quality and maturity level of R&D. *International Journal of Production Economics*, *78*(1), 29-35. doi:10.1016/S0925-5273(00)00166-3

Bhattacherjee, A., & Hikmet, N. (2007). Physicians' resistance toward healthcare information technology: A theoretical model and empirical test. *European Journal of Information Systems*, *16*(6), 725-737. doi:10.1057/palgrave.ejis.3000717

Bozan, K., Parker, K., & Davey, B. (2016). A closer look at the social influence construct in the UTAUT Model: An institutional theory based approach to investigate health IT adoption patterns of the elderly. *Proceedings of the Annual Hawaii International Conference on System Sciences*, *2016-March* (pp. 3105-3114). doi:10.1109/HICSS.2016.391

Brown, S. A., & Venkatesh, V. (2005). Model of adoption of technology in households: A baseline model test and extension incorporating household life cycle. *MIS Quarterly, 29*(3), 399-426. doi:10.2307/25148690

Brünink, L. (2016). *Cross-functional Big Data integration: Applying the UTAUT model*. University of Twente (The Netherlands).

Burton-Jones. (2009). Minimizing method bias through programmatic research. *MIS Quarterly*, *33*(3), 445-471. doi:10.2307/20650304

Cabrera-Sánchez, J.-P., & Villarejo-Ramos, Á.-F. (2018). Factores que afectan a la adopción del Big Data como instrumento de marketing en las empresas españolas. In *XXVIII Jornadas Luso-Espanholas de Gestâo Científica,* At Guarda (Portugal).

Chauhan, S., & Jaiswal, M. (2016). Determinants of acceptance of ERP software training in business schools: Empirical investigation using UTAUT model. *International Journal of Management Education*, *14*(3), 248-262. doi:10.1016/j.ijme.2016.05.005

Chiang, R. H. L., Grover, V., Liang, T.-P., & Zhang, D. (2018). Special issue: strategic value of Big Data and business analytics. *Journal of Management Information Systems*, *35*(2), 383-387. doi:10.1080/07421222.2018.1451950

Chin, W. W., & Dibbern, J. (2010). An introduction to a permutation based procedure for multi-group PLS analysis: Results of tests of differences on simulated data and a cross-cultural analysis of the sourcing of information system services between Germany and the USA. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares* (pp. 171-193). doi:10.1007/978-3-540-32827-8

Claudy, M. C., Garcia, R., & O'Driscoll, A. (2015). Consumer resistance to innovation: A behavioral reasoning perspective. *Journal of the Academy of Marketing Science*, *43*(4), 528-544. doi:10.1007/s11747-014-0399-0

Davis, F. (1985). *A technology acceptance model for empirically testing new end-user information systems*. Massachusetts Institute of Technology, Sloan School of Management (December).

Demoulin, N. T. M., & Coussement, K. (2018). Acceptance of text-mining systems: The signaling role of information quality. *Information & Management*. Advanced online publication. doi:10.1016/j.im.2018.10.006

Ducange, P., Pecori, R., & Mezzina, P. (2018). A glimpse on big data analytics in the framework of marketing strategies. *Soft Computing*, *22*(1), 325-342. doi:10.1007/s00500-017-2536-4

Duyck, P., Pynoo, B., Devolder, P., Voet, T., Adang, L., Ovaere, D., & Vercruysse, J. (2010). Monitoring the PACS implementation process in a large university hospital-discrepancies between radiologists and physicians. *Journal of Digital Imaging*, *23*(1), 73-80. doi:10.1007/s10278-008-9163-7

Escobar-Rodríguez, T., & Carvajal-Trujillo, E. (2014). Online purchasing tickets for low cost carriers: An application of the unified theory of acceptance and use of technology (UTAUT) model. *Tourism Management*, *43*, 70-88. doi:10.1016/j.tourman.2014.01.017

Falk, R. F., & Miller, N. B. (1992). *A primer for soft modeling*. Akron, OH: University of Akron Press.

Featherman, M. S., & Pavlou, P. A. (2003). Predicting e-services adoption: A perceived risk facets perspective. *International Journal of Human Computer Studies*, *59*(4), 451-474. doi:10.1016/S1071-5819(03)00111-3

Fishbein, M., & Ajzen, I. (1975). *Belief attitude, intention and behavior. An introduction to theory and research*. Reading, MA: Addison-Wesley.

Gargallo López, B., Suárez Rodríguez, J., & Almerich Cerveró, G. (2006). La influencia de las actitudes de los profesores en el uso de las nuevas tecnologías. Revista Espanola de Pedagogia, *64*(233), 45-66.

Gefen, D., Rigdon, E. E., & Straub, D. (2011). An update and extension to SEM guidelines for administrative and social science research. *MIS Quarterly*, *35*(2), 3-14. doi:10.1016/j.lrp.2013.01.001

Gibson, C.F. (2004). IT-enabled business change: An approach to understanding and managing risk. *MIT Sloan Working Paper* No. 4520-04; CISR Working Paper No. 346. doi:10.2139/ssrn.644922

Gupta, V. K., Huang, R., & Niranjan, S. (2010). A longitudinal examination of the relationship between team leadership and performance. *Journal of Leadership & Organizational Studies*, *17*(4), 335-350. doi:10.1177/1548051809359184

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, *40*(3), 414-433. doi:10.1007/s11747-011-0261-6

Henseler, J., Ringle, C. M., & Sarstedt, M. (2014). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115-135. doi:10.1007/s11747-014-0403-8

Hsieh, P. J. (2015). Healthcare professionals' use of health clouds: Integrating technology acceptance and status quo bias perspectives. *International Journal of Medical Informatics*, *84*(7), 512-523. doi:10.1016/j.ijmedinf.2015.03.004

Huang, T. C. K., Liu, C. C., & Chang, D. C. (2012). An empirical investigation of factors influencing the adoption of data mining tools. *International Journal of Information Management*, *32*(3), 257-270. doi:10.1016/j.ijinfomgt.2011.11.006

Hung, Y. H., Wang, Y. S., & Chou, S. C. T. (2007). User acceptance of e-government services. *PACIS 2007 Proceedings*, 97. Retrieved from http://aisel.aisnet.org/pacis2007

Ives, B., & Olson, M. H. (2008). User involvement and MIS success: A review of research. *Management Science*, *30*(5), 586-603. doi:/10.1287/mnsc.30.5.586

Khatibian, N., Hasan gholoi pour, T., & Abedi Jafari, H. (2010). Measurement of knowledge management maturity level within organizations. *Business Strategy Series*, *11*(1), 54-70. doi:10.1108/17515631011013113

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision Support Systems*, *44*(2), 544-564. doi:10.1016/j.dss.2007.07.001

Kim, H.-W., & Kankanhalli, A. (2009). Investigating user resistance to information systems implementation: A status quo bias perspective. *MIS Quarterly*, *33*(3), 567-582. doi:10.2307/20650309

Kim, H. W., Chan, H. C., & Gupta, S. (2007). Value-based adoption of mobile internet: An empirical investigation. *Decision Support Systems*, *43*(1), 111-126. doi:10.1016/j.dss.2005.05.009

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of E-Collaboration*, *11*(4), 1-10. doi:10.4018/ijec.2015100101

Kock, N., & Lynn, G. S. (2012). Lateral collinearity and misleading results in variance-based SEM : An illustration and recommendations. *Journal of the Association for Information Systems*, *13*(7), 546-580.

 CC BY

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, *34*(3), 387-394. / doi:10.1016/j.ijinfomgt.2014.02.002

Lai, Y., Sun, H., & Ren, J. (2017). Understanding the determinants of big data analytics (BDA) adoption in logistics and supply chain management: An empirical investigation. *International Journal of Logistics Management*, *29*(2), 676-703. doi. 10.1108/IJLM-06-2017-0153

Lapointe, L., & Rivard, S. (2007). A triple take on information system implementation. *Organization Science*, *18*(1), 89-107. doi:10.1287/orsc.1060.0225

Laumer, S., Maier, C., Eckhardt, A., & Weitzel, T. (2016). User personality and resistance to mandatory information systems in organizations: A theoretical model and empirical test of dispositional resistance to change. *Journal of Information Technology*, *31*(1), 67-82. doi:10.1057/jit.2015.17

Lee, J. H., & Song, C. H. (2013). Effects of trust and perceived risk on user acceptance of a new technology service. *Social Behavior and Personality: An International Journal*, *41*(4), 587-597. doi:10.2224/sbp.2013.41.4.587

Liu, X., Shin, H., & Burns, A. C. (2019). Examining the impact of luxury brand's social media marketing on customer engagement: Using big data analytics and natural language processing. *Journal of Business Research*. Advanced online publication. doi: 10.1016/j.jbusres.2019.04.042

Martins, C., Oliveira, T., & Popovič, A. (2014). Understanding the internet banking adoption: A unified theory of acceptance and use of technology and perceived risk application. *International Journal of Information Management*, *34*(1), 1-13. doi:10.1016/j.ijinfomgt.2013.06.002

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The management revolution. *Harvard Business Review*, *90*(10), 61-68. https://doi.org/10.1007/s12599-013-0249-5

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, *35*(2), 293-334. doi:10.2307/23044045

Norzaidi, M. D., Salwani, M. I., Chong, S. C., & Rafidah, K. (2008). A study of intranet usage and resistance in Malaysia's port industry. *Journal of Computer Information Systems*, *49*(1), 37-47. doi:10.1080/08874417.2008.11645304

Nunnally, J. C. (1978). *Psychometric theory* (2nd edit.) Hillsdale, NJ: McGraw-hill.

Palmatier, R. W., & Martin, K. D. (2019). Understanding and valuing customer data. In R. W. Palmatier & K. D. Martin, *The Intelligent Marketer's Guide to Data Privacy* (pp. 73-92). Palgrave Macmillan, Cham. doi:10.1007/978-3-030-03724-6

Paulk, M. C., Curtis, B., Chrissis, M. B., & Weber, C. V. (1993). Capability maturity model, version 1.1. *IEEE Software*, *10*(4), 18-27. doi:10.1109/52.219617

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879-903. doi:10.1037/0021-9010.88.5.879

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539-–569. doi:10.1146/annurev-psych-120710-100452

Poon, E. G., Blumenthal, D., Jaggi, T., Honour, M. M., Bates, D. W., & Kaushal, R. (2004). Overcoming barriers to adopting and implementing computerized physician order entry systems in U.S. hospitals. *Health Affairs*, *23*(4), 184-190. doi:10.1377/hlthaff.23.4.184

Rahman, N. (2016). Factors affecting Big Data technology adoption, 0–29. *Student Research Symposium*. Retrieved from http://pdxscholar.library.pdx.edu/studentsymposium%5Cnhttp://pdxscholar.library.pdx.edu/studentsymposium/2016/Presentations/10

Rehman, M. H. U., Chang, V., Batool, A., & Wah, T. Y. (2016). Big data reduction framework for value creation in sustainable enterprises. *International Journal of Information Management*, *36*(6), 917-928. doi:10.1016/j.ijinfomgt.2016.05.013

Ringle, C. M., Wende, S., & Becker, J. M. (2015). SmartPLS. "SmartPLS 3." Boenningstedt: SmartPLS GmbH. Retrieved from http://www.smartpls.com

Roldán, J. L., & Sánchez-Franco, M. J. (2012). Variance-based structural equation modeling: Guidelines for using partial least squares in information systems research. In M. Mora, O. Gelman, A.L. Steenkamp & M. Raisinghnani, *Research methodologies, innovations and philosophies in software systems engineering and information systems*. IGI-Global. doi:10.4018/978-1-4666-0179-6.ch010

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, *70*, 263–286. https://doi.org/10.1016/j.jbusres.2016.08.001

Straub, D., Boudreau, M., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information Systems*, *13*(24), 380-427. doi:10.17705/1CAIS.01324

Sun, Z., Sun, L., & Strang, K. (2018). Big Data analytics services for enhancing business intelligence. *Journal of Computer Information Systems*, 58(2), 162-169. doi:10.1080/08874417.2016.1220239

Urwiler, R., & Frolick, M. N. (2008). The IT value hierarchy: Using Maslow's hierarchy of needs as a metaphor for gauging the maturity level of information technology use within competitive organizations. *Information Systems Management*, *25*(1), 83-88. doi:10.1080/10580530701777206

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, *39* (2), 273-315. doi:10.1111/j.1540-5915.2008.00192.x

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*(2), 186-204. doi:10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425-478. doi:10.2307/30036540

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157-178. doi:10.1017/CBO9781107415324.004

Verma, S., Bhattacharyya, S. S., & Kumar, S. (2018). An extension of the technology acceptance model in the big data analytics system implementation environment. *Information Processing and Management*, *54*(5), 791-806. doi:10.1016/j.ipm.2018.01.004

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, *80*(November), 97-121. doi:10.1509/jm.15.0413

Watson, H. J. (2019). Update tutorial: Big Data analytics: Concepts, technology, and applications. *Communications of the Association for Information Systems*, *44*(1), 364–379. doi:10.17705/1CAIS.04421

Wright, L. T., Robin, R., Stone, M., & Aravopoulou, D. E. (2019). Adoption of Big Data technology for innovation in B2B marketing. *Journal of Business-to-Business Marketing*. Advanced online publication. doi: 10.1080/1051712X.2019.1611082

Wu, Y. L., Tao, Y. H., & Yang, P. C. (2007). Using UTAUT to explore the behavior of 3G mobile communication users. *IEEM 2007: 2007 IEEE International Conference on Industrial Engineering and Engineering Management* (pp. 199-203). doi:10.1109/IEEM.2007.4419179

Yadegaridehkordi, E., Hourmand, M., Nilashi, M., Shuib, L., Ahani, A., & Ibrahim, O. (2018). Influence of big data adoption on manufacturing companies' performance: An integrated DEMATEL-ANFIS approach. *Technological Forecasting and Social Change*, 137(March), 199-210. doi:10.1016/j.techfore.2018.07.043

Yaqoob, I., Hashem, I. A. T., Gani, A., Mokhtar, S., Ahmed, E., Anuar, N. B., & Vasilakos, A. V. (2016). Big data: From beginning to future. *International Journal of Information Management*, *36*(6), 1231-1247. doi:10.1016/j.ijinfomgt.2016.07.009

Yu, C.-S. (2012). Factors affecting individuals to adopt mobile banking: Empirical evidence from the UTAUT model. *Journal of Electronic Commerce Research*, 13(2), 104–121.

Zicari, R. (2014). Big data: Challenges and opportunities. In R. Akerkar (Ed.), *Big Data Computing* (pp. 103-128). Boca Raton, FL: CRC Press. doi:10.1201/b16014-5

 CC BY ISSN 0034-7590; eISSN 2178-938X

PERSPECTIVES

Invited article
Original version
DOI: http://dx.doi.org/10.1590/S0034-759020190608

BIG DATA AND DISRUPTIONS IN BUSINESS MODELS

THE HYPE OF EMERGING TECHNOLOGIES

There are many challenges to reaping the benefits of the newest, emerging technologies. If it were easy, every business would do it and competitive advantage would easily fade away. It is in fact extremely difficult and challenging for companies to create value with emerging technologies. Every year, Gartner, a respected consultancy company, reviews the newest, emerging technologies and uses its hype cycle concept to explain the path that technologies take. The cycle consists of five phases: (i) the technology trigger phase: the invention of new technology that happens in a research lab, usually at a university (most companies outsource their fundamental research to universities nowadays). (ii) the peak of inflated expectations phase: the technology is discussed by companies at conferences and in the press. There is a great deal of talk about the new technology, but no one has used it yet. R&D projects are launched. (iii) the disillusionment phase: it turns out that the technology is not as useful as it was thought to be. (iv) the slope of enlightenment phase: here, the valuable fusion of business and technology is explored. (v) the plateau of productivity phase: it is clear how business can use the technology to create value. In Gartner's (2019) emerging technologies hype cycle, technologies such as biorobots, augmented reality cloud, decentralized web, adaptive machine learning, nanoscale 3D printing, and 5G are reviewed and positioned in the first phases. Gartner expect that these technologies will reach the plateau of productivity within 5 to 10 years.

Two observations are important and have to be made. First, the newest technologies will generate enormous amounts of data, both at the level of molecules, individual objects and humans; and interconnected devices (things) and humans embedded in complex commercial, logistical, and financial networks. Second, for companies, it would be wise to address these emerging technologies from a 'sensing—data storage—analyzing and responding—learning' perspective. From a business perspective, the sense-store-analyze-respond-and-learn cycle needs to be connected to the hype cycle of emerging technologies. Completing the loop of 'sensing—data storage—analyzing and responding—learning' is a critical success factor for companies. However, as usual, there are complications, in this case there are two. The first is that more data also will imply more noise. It is very challenging to distinguish the signal from the noise, see Silver (2015). The second complication is that data is generated not only inside the firm's applications but also in applications that are run in conjunction with several business partners in so-called business ecosystems; Porter and Heppelmann (2014) would call these business applications a "system of systems." Therefore, the ownership and legality of data becomes a crucial factor that might hamper technologies from reaching the plateau of productivity.

ERIC VAN HECK[1]
evanheck@rsm.nl
ORCID: 0000-0002-8058-6640

[1]Erasmus University, Rotterdam School of Management, Rotterdam, Netherlands

 CC BY ISSN 0034-7590; eISSN 2178-938X

HOW TO WIN THE DATA RACE?

Many companies are aware of the potential value of emerging technologies and the digital transformation that they need to make. However, as stated previously, it is not easy to transform companies into digital businesses. The MIT Center for Information Systems Research (CISR) has a long research tradition in analyzing and explaining the so-called critical success factors, a term that was coined in the 1980s by Jack Rockart, the first CISR research director. Their latest research shows that it is not so much the technology, but more the organizational capabilities, that drive success. In their book *Leading Digital: Turning Technology into Business Transformation,* Westerman, Bonnet, and McAfee (2014) analyze the performance of 400 companies around the world. They distinguish four clusters (beginners, fashionistas, conservatives, and digital masters) that are distinguished by the capabilities that they incorporate. Digital leadership is about the development of capabilities that will enable the design and implementation of useful business applications (WHAT) and those that will enable transformation (HOW). The first group, digital capabilities, enables the creation of a compelling customer experience, exploiting the power of core operations, and reinventing business models. The second group, leadership capabilities, deals with crafting a digital vision, engaging the organization at scale, governing the transformation, and building technology leadership capabilities. Both digital and leadership capabilities are necessary conditions for attaining higher firm performance levels. Companies that incorporate these capabilities perform exceptionally better, both in terms of revenue generation, profitability, and market valuation. Digital transformation programs are crucial for the survival of companies. Take, for example, Royal Philips, an established, century-old conglomerate with dwindling results. With a new CEO in 2011, it established a digital transformation program, called Accelerate, that created a new foundation for a digitally enabled company that moved from product to software service orientation, see Mocker, Ross, and van Heck (2014). A crucial component of the transformation was to eliminate unrewarded process complexity and simplify processes, roles, and systems around the world. By contrast, the transformation stimulated rewarding product complexity such as integrated health solutions with advanced scanning, decision analysis, and artificial intelligence components. Recently, Peter Weill and Stefanie Woerner (2018) published the latest CISR research in their book *What's Your Digital Business Model? Six Questions to Help You Build the Next-Generation Enterprise*. They emphasize that digital transformation is not about technology but rather about change. They distinguish four business models based on business design dimensions (value chain versus ecosystem) and end customer knowledge (partial versus complete), namely: supplier, omnichannel, modular producer, and ecosystem driver. Ecosystem drivers outperform all other business models.

WHAT IS THE BUSINESS VALUE AND FOR WHOM?

There are three fundamental ways for companies to create business value:

- To create business that is more efficient—resources are used in a better way
- To create business that is more effective—end customers are served in a better way
- To create new business—new business models create sustainable value.

In all three of these value creation strategies, the use of Big Data and advanced analytics play a crucial role. First, take, for example, the retailer's efficiency of the last mile delivery of packages that customers purchase online. Albert Heijn, a large retailer in Europe, felt a need to redesign its last mile delivery process. It came up with an intelligent approach that combined the data of the incoming online calls for delivery, with the pricing of time slots for delivery, and the efficient bundling of deliveries in trucks that served the same area (Agatz, Campbell, Fleischmann, & Savelsbergh, 2011). The result was a higher utilization of truck capacity and a lower number of truck kilometers reducing the carbon footprint of last mile deliveries in cities.

Second, Royal FloraHolland, the world leading market grower of flowers and plants, is transforming to become an ecosystem driver that is more effective in the floriculture world. One way is to redesign the flower auction markets and augment the auctioneers with advanced analytics capabilities. Its customers, in this case professional wholesalers and retailers, will get the right products for the right price at the right time. Because of the perishability of flowers, time is a crucial source of value. Based on the analysis of millions of transactions, they created an advanced decision support system that enable auctioneers to make better decisions and balance revenue with turnaround time (Lu, Gupta, Ketter, & van Heck, 2019).

Third, the most challenging strategy is to create new business. Car2Go, an internationally operating car rental and sharing company, rents out cars to customers. In cities such as

 CC BY © RAE | São Paulo | **59(6)** | **November-December 2019** | 430-432 ISSN 0034-7590; eISSN 2178-938X

New York, Amsterdam, and Stuttgart, one can pick up a car from the street and leave it anywhere. A pilot with electric cars in San Diego, Amsterdam, and Stuttgart showed the potential of a new business model for electric cars. With the help of advanced algorithms and different data sets, such as weather, short-term local electricity prices, and car rental demand, the fleet of electric cars can simulate a virtual power plant (Kahlen, Ketter, & van Dalen, 2018). The combined batteries of these electric cars can store electricity generated by the sun and wind, the renewable energy providers. On overcast and/or still days, electricity stored in the electric car batteries can be sold via the smart grid to the general public. Therefore, Car2Go might shift its business model from a "car rental" to a "car rental and electricity storage and sales" company.

WHAT ARE THE TRANSFORMATIONAL CHALLENGES?

Digital transformation has many challenges and nowadays most companies face these challenges. The first is that most companies do not have a common digitized platform. They have a plethora of applications and systems that were built with relatively outdated technology and software, such as Enterprise Resource Planning applications. The transformation to real time access to enormous data sets with advanced human and machine learning capabilities will take some time and effort.

The second challenge is how employees are able to work with the new advanced machine intelligence capabilities. Are employees engaged in the transformation or do they feel threatened that robots will take over their jobs? In most organizations, users of technology cannot or will not engage in the transformation.

The third challenge is the lack of digital leadership by the board of the company. Most board members are not digital savvy and therefore it is difficult for them to create a vision and mission for the company that is sustainable in a digitizing world. The board's lack of commitment to transformation is usually a sign to the rest of the company to slow down the pace of change. The fourth challenge relates to data ethics, privacy, and algorithm accountability. Companies acquire detailed data about their customer preferences. Digital technology moguls, such as Alibaba, Amazon, Apple, Facebook, Google, Microsoft, Samsung, and Tencent gather very detailed data about every aspect of their customer's life cycle. Combined with advanced machine learning tools, these databases create data power and stimulate an important discussion about the rules and responsibilities of companies. Who is responsible when companies create algorithms that discriminate against employees or customers? Who owns the data that is generated by customers? In May 2018, the European Union implemented a new regulation called the General Data Protection Regulation, that provides data protection and privacy for all citizens of the European Union and the European Economic Area, including the transfer of personal data outside the European space. In other areas around the world, most citizens are not protected at all.

REFERENCES

Agatz, N., Campbell, A., Fleischmann, M., & Savelsbergh, M. (2011). Time slot management in attended home delivery. *Transportation Science, 45*(3), 435-449. doi:10.1287/trsc.1100.0346

Gartner. (2019). *Hype cycle of emerging technologies 2019*. Retrieved from https://www.gartner.com/en/documents/3956015/hype-cycle-for-emerging-technologies-2019

Kahlen, M. T., Ketter, W., & van Dalen, J. (2018). Electric vehicle virtual power plant dilemma: Grid balancing versus customer mobility. *Production and Operations Management, 27*(11), 2054-2070. doi:10.1111/poms.12876

Lu, Y., Gupta, A., Ketter, W., & van Heck, E. (2019). Dynamic decision making in sequential business-to-business auctions: A structural econometric approach. *Management Science, 65*(8), 3853-3876. doi:10.1287/mnsc.2018.3118

Mocker, M., Ross, J. W., & van Heck, E. (2014). Transforming Royal Philips: Seeking Local Relevance While Leveraging Global Scale. In *MIT Sloan CISR Working Paper No. 394, February 2014*. Cambridge, USA: MIT Sloan Centre for Information Systems of Research (CISR).

Porter, M. E., & Heppelmann, J. E. (2014). How smart, connected products are transforming competition. *Harvard Business Review*. Retrieved from https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition

Silver, N. (2015). The signal and the noise: Why so many predictions fail, but some don't. New York, NY: Penguin Books.

Weill, P., & Woerner, S. (2018). *What's your digital business model? Six Questions to help you build the next-generation enterprise*. Cambridge, UK: Harvard Business Review Press.

Westerman, G. Bonnet, D., & McAfee, A. (2014). *Leading digital: Turning technology into business transformation*. Cambridge, UK: Harvard Business Review Press.

 CC BY ISSN 0034-7590; eISSN 2178-938X

PERSPECTIVES

Invited article
Original version
DOI: http://dx.doi.org/10.1590/S0034-759020190609

PLUS ÇA CHANGE, PLUS C'EST LA MÊME CHOSE

In 1572, a single data point, Tycho's Supernova, showed that contrary to the accepted paradigm at the time, the heavens did indeed change (Wootton , 2015). Less than 40 years later, in 1610, Galileo Galilei published his sensational findings in *Sidereus Nuncius*, a short treatise that demonstrated the existence of stars not seen by the naked eye and revealed the nature of the Milky Way (Galilei, 1610). Since then, data analysis has been central to scientific research and examples of its use in solving important and difficult problems have multiplied. At the age of 24, Carl Friedrich Gauss (1809) used least squares to correctly predict Ceres' position in 1801 after it emerged from behind the Sun's glare. A simple spatial analysis identified the source of the Broad Street cholera outbreak in London in 1854 (Snow, 1855). Between 1856 and 1863, careful estimation of frequencies allowed Gregor Mendel (1866) to determine the basic rules of heredity of physical traits in plants. In the late 1940's, a large retrospective study led by Richard Doll and A. Bradford Hill (1950) demonstrated the strong link between smoking and lung cancer.

The steady development and refinement of new statistical methods after 1880 by Galton, Student, Fisher, and others allowed for a broader range of applications of data analysis methods in industry and business. Ideas and methods developed and popularized by Shewhart, Deming, and others made statistical quality control an integral part of the industrial manufacturing process. This also incorporated the use of modern experiment design after the war. Inexpensive computing power, automated data collection, and the development of some general and flexible data analysis software—especially *R*, which is both comprehensive and free—greatly expanded the spectrum of applications. Consequentially, the era of "Big Data" was born. If much of what is published in the press is to be believed, Big Data will solve critical problems in areas as disparate as medical diagnostics, credit evaluation, weather forecast, and facial recognition. We will have much better products and services along with a much deeper understanding of physical and cultural processes as a result of ever larger data sets and the modern computer's intensive methods.

While making predictions is a difficult business, I am sure the key issues we will face in applying statistical analysis methods to business problems in the era of Big Data will continue to be the same ones we have been dealing with for decades. Data analysis still is about information, insights, and conclusions. Data is used to tell a story; analytical thinking is required in the evaluation of the relationship between data and conclusions. In regard to that evaluation, the three big challenges, in order of importance, have been and will continue to be overstating statistical significance; lack of reproducibility; and the avoidance of providing exact answers, but to the wrong questions.

Statistical significance is viewed by many as the gold standard. Significance, simply put, is the probability that an observed set of observations is the result of random fluctuation. If the calculated value is small—usually less than 5%—one is led to believe that some factors must exist

FLAVIO BARTMANN[1]
fcb2122@columbia.edu
ORCID: 0000-0002-9308-3049

[1]Columbia University, School of International and Public Affairs, New York, NY, United States of America

 CC BY

that explain the non-random behavior. Behind the calculation of the *p*-value (the usual measure of significance), there is always a complex process of study design, data collection, and data analysis, including model selection. The level of significance claimed is only accurate if all the elements of the process are done well and if the model chosen provides a reasonable description of reality. However, this is rarely the case. The situation is particularly dire in the business context where the models used can be very complex. These intricate models are typically regression models with several explanatory variables where the relationships with the response are assumed to be linear and the calculation of the estimated coefficients' significance is frequently meaningless. Poor design and flawed data collection only add to the troubles. The problem is so serious that a large movement to demote or even eliminate p-values has gathered much support among statisticians (McShane, Gal, Gelman, Robert, and Tackett, 2019).

The second big problem with data analyses is that many—perhaps a majority—are not reproducible (Ioannidis, 2005). The most common cause is one or more flaws in the design of the investigation, be it an experiment or a survey. Flaws in the data collection process, use of inappropriate methods, and fraud are also common occurrences that invalidate studies. In scientific research, studies can be re-done and some degree of self-correction can be achieved. The problem is much more serious within industry environments where business decisions are frequently data-driven and time sensitive, making replication very rare. A poorly carried out market or feasibility study can have costly consequences. Remember Apple's Newton platform?

The third problem might be the most serious and the most difficult to address. John Tukey (1962) used to say that "an approximate answer to the right question is worth a great deal more than a precise answer to the wrong question." The classical story, told numerous times by statisticians everywhere (or "data scientists" in modern parlance), but even more frequently in the corridors of the Department of Statistics at Columbia University, is about the reinforcement of British planes used in the bombings of Germany late in the Second World War. A large proportion of the planes were lost due to the German anti-aircraft fire and accordingly the decision was made to bolster them with armor. The most important concern to address for those on this project was which location was the best place to put the additional armor. The returning planes were carefully examined, and a proposal was made to put the extra armor in the areas that had received the most damage by fire. Interestingly, it would have proved to be a fatal mistake if this decision had been carried out. The correct variable that should have been reviewed was not regarding the planes that could be examined, but about the planes that had not returned. Fortunately, Abraham Wald cleverly suggested that perhaps areas where the returning planes were unscathed were those that should have been reinforced.

Data sets, even those with terabytes of records, are merely the raw material of knowledge. Today, almost everything can be monitored and measured but the key challenge continues to be our ability to use and analyze these data sets and to make sense of them in order to tell the real story.

REFERENCES

Doll, R., & Bradford Hill, A. (1950) Smoking and Carcinoma of the Lung. *British Medical Journal*, *2*(4682), 739–-748. doi:10.1136/bmj.2.4682.739

Gauss, C. F. (1809). *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. Hamburg: Friedrich Perthes and I.H. Besser.

Galilei, G. (1610). *Sidereus Nuncius*. Thomam Baglionum: Venice.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLOS Medicine*, *2*(8), e124. doi:10.1371/journal.pmed.0020124

McShane, B. B., Gal, D., Gelman, A., Robert C., & Tackett, J. L. (2019). Abandon statistical significance. *The American Statistician*, *73*, 235–245. doi:10.1080/00031305.2018.1527253

Mendel, G. (1866). Versuche über Plflanzenhybriden. Verhandlungen des naturforschenden Vereines in Brünn, Bd. IV für das Jahr 1865 (pp. 3-47). *Abhandlungen*.

Snow, J. (1855). *On the Mode of communication of cholera*. London, UK: John Churchill.

Tukey, J. W. (1962). The future of data analysis. *Annals of Mathematical Statistics, 33*, 1-67. doi:10.1214/aoms/1177704711

Wootton, D. (2015). *The invention of science: A new history of the scientific revolution*. London, UK: Allen Lane.

 CC BY © RAE | São Paulo | **59(6)** | **November-December 2019** | 433-434 ISSN 0034-7590; eISSN 2178-938X

ESSAY

Submitted 03.13.2019. Approved 08.19.2019
This article is authorship by a member of RAE's Scientific Editorial Board and was evaluated by double blind review process with impartiality and independence. Scientific Editor: Marcelo Bispo
Translated version
DOI: http://dx.doi.org/10.1590/S0034-759020190610

CORPORATE CRIMES: THE SPECTER OF GENOCIDE HAUNTS THE WORLD

Exploitation of workers, communities and resources has always occurred, although it is not a central topic in the history of management, whose main narrative is urgency in the pursuit of efficiency and the right solutions. However, crimes, misconduct, unethical behavior and corporate social irresponsibility have become increasingly common, thus calling for reflections on the relevance of addressing these issues in the academic field and in management practice. In this essay, we start from the premise that corporate crime must be understood according to its multidisciplinary nature, and we focus our efforts on discussing management issues to argue that corporate crime is part of companies' operations to support contemporary capitalism. We begin by addressing the power of corporations as the main force of contemporary capitalism in its form of extreme concentration of corporate wealth and ownership. Then we discuss the seriousness of corporate crimes and how they resemble genocide. We close with our considerations on why organizations become criminal and present a way to prevent corporate crime.

CORPORATIONS

The Industrial Revolution gave a new outline to the operation of companies, especially with the emergence of new forms of business organization, such as the modern corporation, whose distinctive feature is the separation between ownership, which is scattered among many shareholders, and the control exercised by directors who own a small fraction, if any, of the company's stocks (Berle & Means, 1932).

The emergence of corporations in the nineteenth century changed the ownership mechanisms of companies, including in legal terms, and rapidly, from 1840 to 1860, corporations became capitalists' preferred business organization model. In the 1870's, the major corporations in most industries sought to reduce competition and increase their profits through a horizontal combination that allowed them to control raw materials and the market, as well as other advantages. The main characteristic of this type of business organization is its capacity to exert influence and power over a large geographical area, including in cultural and social terms, in addition to the possibilities of making higher profits. By then, corporations were already being accused of fixing prices, exploitation and other abuses, thus resulting in a political reaction by the US government, which created new forms of regulation for this type of organization in the late 19th century (Clinard, Yager, Brissette, Petrashek, & Harries, 1979).

CINTIA RODRIGUES DE OLIVEIRA[1]
cintia@ufu.br
ORCID: 0000-0001-7999-9002

[1]Universidade Federal Uberlândia, Faculdade de Gestão e Negócios, Uberlândia, MG, Brazil

 CC BY

The effects of the changes arising from the consolidation of corporations divide researchers' opinions. Advocates for the positive effects point to the development of new goods and services, lower prices, job creation, improved formal education of people, and prosperity for the lower-income classes. In turn, those who point out the negative effects criticize the influence that corporations have on people's lives, especially due to their ideological power, which legitimizes their *status quo* as the only, desirable one. Large businesses imply high economic power in the hands of few; employees have a limited bargaining power with their employers and are more vulnerable to economic downturns, in addition to being forced to accept precarious working conditions with long working hours and low wages (Pearce & Tombs, 1999).

Considering the social and economic influences, one cannot forget that the state, in part of the world, has abandoned its intervener role: state-owned companies were privatized, new financial and fiscal instruments facilitated more efficient forms of production, and free-trade policies on a global level brought about a globalized system that has undermined the negotiating power of regions and nations (Bauman, 1999). It is no exaggeration to say that the state has become an apparatus used by a power bloc, e.g., corporations in a particular industry, to secure, maintain and expand its dominance over the economy, politics and society, as well as regulatory agencies (Pearce & Tombs, 1999). Corporations thus acquired the power of mobility, which allowed them to choose the resources, labor and location of their operations so as to obtain advantageous production conditions while eliminating any kind of limitations and constraints.

As for the relationship between governments and corporations, the latter make use of political connections, such as donations to political campaigns and the inclusion of members with a political background into boards of directors (Camilo, Marcon, & Bandeira-de-Mello, 2012). In analyzing the American context, Barley (2007) shows the power of influence of corporations on social institutions, including in democracies, contradicting traditional assumptions of organizational theorists that only the external environment affects organizations and they affect each other. Barley (2007) maps lobbying connections in the United States, representing corporations with US politicians to intervene in the environment in order to secure corporate interests by influencing social institutions in three ways: (1) creating legislation that benefits corporate citizens ; (2) limiting the creation of regulatory agencies that protect the public good from corporations' acts and the externalities they create; and (3) privatizing roles that should be performed by local, state and federal governments.

And so the state lost strength in its mediating role between the market and society, thus giving rise to a "new proliferation of weak and powerless sovereign states" (Bauman, 1999, p. 75), while multinationals consolidated their influence and the power to obtain concessions and settle in a business environment fostered by the abundance of low-cost, skilled labor, as well as the low regulation of working conditions. This influence and power grow as corporate forms of organization change, such as a general wave of mergers that began in the latter half of the twentieth century, particularly in the United Kingdom and the United States. Mergers, acquisitions, strategic alliances and joint ventures have become common strategies for corporations, since they allow them to share costs and risks as they increase their profits, market and power, and make politicians powerless to exercise any control over them.

In their struggle for survival in an environment of fierce competition for resources, corporations, whether intentionally or locally, indirectly or directly, conduct themselves in ways that can lead to crimes that are often dragged to the backstage of social life. Such conduct, decisions and actions within large corporations may constitute crimes against society, consumers, employees, the community, investors, governments and the environment. Such business actions came to be termed corporate crimes by journalistic texts and specialist websites, and the term became widely used in the last decade of the twentieth century (Erp, 2018).

Russell Mokhiber, a journalist and editor at the American weekly newsletter *Corporate Crime Reporter*, founded in 1987, compiled high-profile cases of crimes committed by companies in his 1988 book *Corporate Crime and Violence: Big Business Power and the Abuse of the Public Trust*. Mokhiber (1995) profiles 36 cases of corporate crime and violence that occurred until the 1980's, with details of violence against women, children, the environment and consumers, with an emphasis on the damaging consequences of business conduct. Corporate crimes have become increasingly common and are covered by the media under various denominations, being clearly described as a problem that transcends the individual level to lay its roots in corporate structures, processes and decisions (Erp, 2018).

Corporate crimes are widely discussed in sociology, law and economics, but approaches can vary, and their origins are not always associated with criminal conduct. Regarding the conceptualization of the term corporate crime, there is abundant terminology, including the term white collar crime (e. G. Sutherland, 1940), considered one of the first references to crimes committed in suits, in addition to the terms occupational crime and organizational crime.

 CC BY © RAE | São Paulo | **59(6)** | **November-December 2019** | 435-441 ISSN 0034-7590; eISSN 2178-938X

The specialist literature on corporate crime associates it with past business performance, pressures and barriers to achieving superior performance, structure, environment, technology, and other organizational variables. In addition, the principle of autonomy ensures that owners and shareholders will never bear the costs of social harm caused by corporations (Whyte, 2018). Nor do corporations bear the costs of preventing corporate crime from happening; instead, they prefer other practices that transfer these costs to society, which hardly ever associates the word crime with events in the corporate environment, even though they occur in pursuit of corporate profit and performance goals. However, "How to prevent corporate crimes?" is not a simple question to answer. It requires efforts to understand, firstly, the dimension in which they occur in society and, secondly, the interrelationships between the various actors involved in committing them.

CORPORATE CRIME SERIOUSNESS

In the literature on corporate crime, one aspect worth noting is the public opinion on the subject: ordinary crimes (or 'street crimes') are scarier than corporate crimes, even though the latter entail a set of damages that go beyond those entailed by the former, i.e., they reach invaluable proportions, since official statistics on their costs do not take into account diseases caused by environmental pollution, the sale of products that can harm consumers, potentially hazardous working conditions due to exposure to toxic products, among others, which make their costs underestimated. In addition to the contextual difficulties in identifying and measuring the costs and consequences of these crimes, in many of them the real damages are not reported in order to avoid embarrassment to the businesses involved.

Overall, without considering specific violence contexts, the financial costs of white collar crimes are as high as or higher than the costs of crimes considered "ordinary" or street crimes (Cohen, 2015). But the damage to social relations goes beyond the financial loss caused by that kind of crime. This is because white collar crime violates trust, thus generating low social moral and producing large-scale social disorganization, which is not the case with street crime, whose effects on social institutions and organizations are smaller.

Indeed, the high costs of corporate crime exceed by far the costs of individual crime, since in the former a single, simple act can result in thousands of victims. Violation of occupational safety standards can lead to many deaths and accidents; environmental contamination and pollution can affect many families and communities; the use of hazardous materials in the production of goods can increase the risk of health problems for many workers and consumers. In addition, victims are not just individuals, but also small businesses that can be led to insolvency.

The discussion of corporations' liability and culpability for damage caused by their activities comes down to three main points: the first concerns the fact that a corporation can never be arrested; the second is that one must recognize that if corporations are subject to criminal law just like individuals, then this could mean that they have the same rights and responsibilities; and the third refers to the tolerant attitudes or reactions of society to organizations' conduct. This third aspect stems mainly from the exaltation of the market and private enterprise as responsible for the progress and economic development of nations, which has led to a sanctifying stance towards corporations.

Sutherland (1940) argued that the small number of corporate crime convictions in the US criminal justice system was partly due to the absence of effective criminal sanctions for this type of crime. He explains this absence by referring to the impossibility of sentencing a corporation to death or imprisonment, the only possible penalty being fines, which, in fact, are paid by the shareholders in the form of reduced dividends. Sutherland also argues that white collar crime has found room to grow due to public tolerance to it, which has been changing over the years.

It is a fact that public opinion plays a key role in the debate on corporate culpability, given its influence on the control of illegal conduct in the business world. The population in general considers corporate offenses to be serious only when their consequences are physical, substantial, and relatively immediate. However, white collar crime is a real crime, and even if it is not commonly referred to as such, that does not make it smaller. Public opinion also plays a relevant role in regulating and controlling corporate crimes, so much so that the ambiguity of public opinion regarding illegal corporate behavior causes the law to be ambiguous too.

This issue was explored in a survey (Unnever, Benson, & Cullen, 2008) conducted with Americans to find out whether they wished to see stricter stock market regulations enacted and advocate more punishing criminal sanctions for executives who hide the company's true financial condition. As these authors reinforce, the public's feelings are potentially important in shaping crime control policies, especially if they occur when the message from the public is that something must be done to curb corporate crime.

In the US context, since the 1980's, street crime has received increased attention from the government, thus resulting in the adoption of more punishing public control policies against crime. However, with regard to corporate crime, despite the wave

 CC BY © RAE | São Paulo | **59(6)** | **November-December 2019** | 435-441 ISSN 0034-7590; eISSN 2178-938X

of corporate scandals in the country, the subject has not been treated with the necessary attention, which resulted in a gap to be filled. Most corporate violations were dealt with through civil and regulatory procedures rather than by the criminal justice system, which gives these violations the character of an accidental event.

A few reasons justify this treatment of corporate violations of the law. The first of these is the belief among economists and policymakers that the free market is self-regulating and there is no need for criminal law to control harmful corporate behavior. From an opposite viewpoint, authors like Snider (1990) believe that government controls, in the form of criminal regulations or laws, are necessary and appropriate.

The movement against corporate crime in the US has taken place since 1970, resulting in an increase in the use of criminal law against corporations. Unnever et al. (2008) analyze the evolution of criminal legislation and regulation directed at the corporate world, in the United States, in three steps: (1) a particular type of scandal is uncovered and (2) public outcry calls for action by government, which reacts (3) by making a formal accusation or creating new laws and regulations.

Reactions by the population to corporate conduct are relevant to establishing public policies for criminalizing corporate crime, since the legitimacy of a country's criminal justice system can be challenged if the government fails to respond to high-profile corporate crimes that result in damage to a significant number of victims.

The institutions and corporations present in society reflect economic and social inequalities, thus producing transgressive practices and patterns, which in turn result in the economic and political normative processes in this society. Therefore, a critical view of corporate crime can recognize that crime, like social processes, is shaped by those who enjoy political and economic power and influence to ensure that the denomination of what crime is reflects their worldview and interests, in the case of corporations, economic, social and political power.

CAN WE SPEAK OF GENOCIDES?

Corporate criminal actions have gained the status of genocides (Brook, 1998; Kelly, 2013, 2016; Medeiros & Alcadipani, 2018; Stokes & Gabriel, 2010), and organizations and their instrumental rationality have also been associated with the Holocaust (Bauman, 1998; Black, 2001). Despite the different meanings that the term has acquired throughout history, as well as the controversies about it, it is no exaggeration to make such associations because genocide is defined as the mass murder committed in an organized manner, and even though it is done by the State and militias, corporations are accomplices to it (Stel, 2014) in that they provide weapons and other resources. Although mass killings occurred earlier, it was not until the twentieth century that the term genocide was coined by Polish Jewish jurist Raphael Lemkin, in 1943, then a refugee in the United States, in an effort to denounce Nazi atrocities. Undoubtedly, Lemkin was influenced by the Holocaust, which victimized many members of his family, to define, in his book Axis Rule in Occupied Europe (1944), genocide as the murder of ethnic, religious, and national groups (Naimark, 2015).

Apart from corporate crimes that directly cause the death of hundreds of people, financial companies act as accomplices to governments and other organizations that commit genocide, such as BNP Paribas, which is accused of financing the purchase of weapons used in the genocide in Rwanda, in 1994 (BBC NEW, 2014). Another example is the collaboration of IBM and Ford with the Holocaust (Lima, 2016). Complicity to genocide, which can be characterized in different ways, is an act punishable under Article 3 of the Convention on the Prevention and Punishment of the Crime of Genocide, an international document prepared after World War II to protect the human person from Nazi genocides. In the same document, genocide is defined by acts committed to destroy a national, ethnic, racial or religious group, whether in part or as a whole. This concept is controversial because it does not include political, economic and cultural groups, which are deliberately excluded from the definition (Schabas, 2009). It also excludes the lives dismissed by contemporary capitalism, such as the deaths caused by the unbridled pursuit of economic profit, which Banerjee (2008) calls necrocapitalism.

Thus, the destruction caused by corporate crimes also occurs on a large scale as with genocide. The No Business with Genocide campaign was created in 2017 to prevent corporations from doing business with regimes that engage in genocide or crimes against humanity. Mokhiber (1995), in arguing that corporate crime is more violent than street crime, supports his arguments by presenting worldwide and US statistics about people murdered in the streets in the United States and around the world and about those who die in same period due to occupational illnesses and lack of safety at work, as well as victims of products that are harmful to health.

Union Carbide caused more than 3,000 deaths and left more than 50,000 people unable to work, in the case of Bhopal, India, in 1984. The mining industry impacts human rights and the environment, destroying lives. In Brazil, a mining company was convicted to pay a compensation for environmental and social damages as a result of contamination by lead which

 CC BY © RAE | São Paulo | **59(6)** | **November-December 2019** | 435-441 ISSN 0034-7590; eISSN 2178-938X

affected more than 18,000 people in the city of Santo Amaro da Purificacao, in the state of Bahia. Forty-six million people live in contemporary slavery. The war industry continues to supply weapons for conflicts, producing billions of bullets and millions of increasingly sophisticated weapons, with a trillion and a half dollars spent on weaponry, while one person is hit by a bullet each second. And hundreds of garment workers in Bangladesh, Dhaka and Chittagong died in the collapsed buildings where they worked for suppliers to major brands.

One cannot precisely count the number of lives lost or the material and social losses, which are difficult to identify and measure in corporate crime. In addition to all the physical deaths of people, rivers, fauna and flora, we must consider the psychological death of those who survive and the insecurity of the population about information that has surfaced about the fragile operating systems of companies that commit corporate crimes, which was formerly kept secret.

WHY ORGANIZATIONS BECOME CRIMINAL

Criminal behavior within corporations should not be analyzed as a personal deviation, but rather as the product of human relationships and interactions in specific contexts, depending on their economic, legal, organizational and normative characteristics, because regardless of the degree of personal motivation of those involved, corporate crime is a legitimate activity in the context in which it emerges.

It is not simple to say why organizations become criminal. Perhaps this answer will arise from studies with inter and multidisciplinary approaches, in their different perspectives and disciplines. In addition to the findings of previous research on the causes of corporate crime, which point to organizational and institutional factors, organizational culture, pressure for results and economic constraints, among others, we address reflections on how corporate crimes are organized by corporations.

The first of these concerns the participation of government agencies in two ways: when government-contracted corporations engage in deviant practices or have government approval to do so; and when government regulatory institutions fail to restrict deviant business activities. This is the so-called state-corporate crime, a type that comprises the intersection between governments and corporations to produce social harm, and the term was first presented by Richard Kramer in 1990, during the annual meeting of the Society for the Study of Social Problems.

Another reflection concerns the willingness that corporations naturally have to commit crimes or break the law in the interest of maximizing profits (Tombs & Whyte, 2015). Organizations provide opportunities for mobilizing the knowledge needed to commit crimes, for example, by keeping secrets, disguising illegal acts, omitting illegal profits, destroying evidence, paying lawmakers, politicians and authorities not to apply the law to them. It is therefore possible to conclude that organizations are a weapon for committing corporate crimes, and that organizations collaborate with each other to commit corporate crimes, either through joint ventures or other strategic alliances.

One must also consider the conduct of organization leaders who are responsible for decisions that lead to corporate crimes in the pursuit of corporate goals. Thus, motives related to economic factors, the decisions of managers/executives and the relations established with the State interconnect so that organizations commit crimes and subsequently become repeat offenders.

Corporations, then, do not become criminals. If they are the engine of contemporary capitalism, and profit is the main goal, some crime will potentially occur. If indemnities are paid, they do not impact profits. In most cases, even though the stocks of a company that committed corporate crime have dropped, and despite its losses, after a while the company is able to pick up its profits and pay generous bonuses to its executives.

BASES FOR REFLECTIONS ON CORPORATE CRIME CONTROL

This scenario is daunting: the dominance of corporations over our lives makes it impossible to dismantle the potentiality of corporate crimes whose lethality can be compared to that of genocide. We do not have an answer to the simple question "what to do?". Perhaps one thing to bear in mind is that corporate crimes should not gain the status of corporate malpractice and be prosecuted as such, that is, under civil law. In such cases, court decisions result in fines that are paid to the government or, in some cases, compensation to the injured parties. We reject the idea that the State should not intervene by means of regulations, prohibitions and severe punishment to corporations and their leaders. And yes, we agree with intervention that implies the loss of autonomy by this business model, which also means the removal of privileges that exempt owners from all the damage they cause (Whyte, 2018).

Change is needed in the types of political leadership, especially concerning the funding of or donations to political campaigns by corporations and companies. Controlling abuse by corporations requires emancipating politicians from these bonds, the price of which is often high and paid with lives.

 CC BY © RAE | São Paulo | **59(6)** | **November-December 2019** | 435-441 ISSN 0034-7590; eISSN 2178-938X

The connections between corporations and legislators are hindrances to a state of things where we no longer have to mourn corporate crime victims. Almost 30 years ago, Sargent (1990) already warned about the need to problematize the criminalization of corporate violations, while also calling for efforts to be directed at exploring other avenues for controlling corporate crime. Thus, society, represented by its citizens, is expected to question the practices and conduct of corporations that affect its current and future way of life. The social change required to prevent corporate crimes, or rather, to keep us from bearing the brunt and harm caused by them, which are profitable for corporations, has its genesis in questioning the dominance of corporations in contemporary society, as well as their responsibilities for corporate transgression.

The corporation as a business model is the greatest force in contemporary capitalism. The principle of corporate separation, under which corporate responsibilities and asset ownership are exclusively the corporation's, so shareholders have no responsibility for its crimes and misconduct, is an encouragement for corporate decisions to be made without considering their harmful consequences. No society wants to mourn lost lives or claim compensation for its material losses anymore, or shed tears for the destruction of its built identities, or even relive traumatic memories. What society wants is for corporate crimes to no longer happen. And in our view, dissolving this form of organizing business, the corporation, as well as others that facilitate the extreme accumulation of wealth and power, would be a way to prevent corporate crime.

We close this essay with a few points to be considered as a basis for our reflections on corporate crime control, since in our view they cause immeasurable damage and incalculable costs, including lives, which are priceless. It is not a matter of proposing solutions to a complex social problem such as the criminal activity of corporations, but rather providing the basis for an analytical reflection on corporate crimes as genocides.

Acknowledgements

This article is part of the results of research conducted during the postdoctoral stage in Business Administration at Fundação Getulio Vargas, Escola de Administração de Empresas de São Paulo. This study was supported by CNPq.

The authors would like to thank the reviewers for their careful reading and suggestions that greatly contributed to improve the initial version.

REFERENCES

Banerjee, S. B. (2008). Necrocaptalism. *Organization Studies*, *29*(12), 1541-1563. doi: 10.1177/0170840607096386

Barley, S. R. (2007). Corporations, democracy, and the public good. *Journal of Management Inquiry*, *16*(3), 201-215. doi: 10.1177/1056492607305891

Bauman, Z. (1998). *Modernidade e Holocausto*. Rio de Janeiro, RJ: Jorge Zahar Editor.

Bauman, Z. (1999). *As consequências da globalização*. Rio de Janeiro, RJ: Jorge Zahar Ed.

BBC NEWS (2014). Entenda o genocídio de Ruanda de 1994: 800 mil mortes em cem dias. 07/04/2014. Disponível em: https://www.bbc.com/portuguese/noticias/2014/04/140407_ruanda_genocidio_ms. Acesso em: 24 set.2019.

Berle, A., & Means, G. (1932). *The modern corporation and private property*. New York, USA: Macmillan.

Black, E. (2001). *IBM and the Holocaust: The strategic alliance between Nazy Germany and America's most powerfull corporation*. New York, USA: Crown Publisher.

Brook, D. (1998). Environmental genocide: Native Americans and toxic waste. *The American Journal of Economics and Sociology*, *57*(1), 105-113. doi: 10.1111/j.1536-7150.1998.tb03260.x

Camilo, S. P. O., Marcon, R., & Bandeira-De-Mello, R. (2012). Conexões políticas e desempenho: Um estudo das firmas listadas na BM&FBOVESPA. *Revista de Administração Contemporânea*, *16*(6), 806-826. doi: 10.1590/S1415-65552012000600003

Clinard, M. B., Yager, P., Brissette, J., Petrashek, D., & Harries, E.(1979). *Illegal corporate behavior.* Washington D.C., USA: Department of Justice.

Cohen, M. A. (2015). Willingness to pay do reduce white-collar and corporate crime. *Journal of Beneffit-Cost Analysis*, *6*(2), 305-324. doi: 10.1017/bca.2015.43

Erp, J. van. (2018). The organization of corporate crime: Introduction to special issue of administrative sciences. *Administrative Sciences*, *8*(36), 1-12. doi:10.3390/admsci8030036

Kelly, M. J. (2013). "Never again"? German chemical corporation complicity in the Kurdish genocide. *Berkeley Journal of International Law*, *31*(2), 348-384.

Kely, M. J. (2016). *Prosecuting corporations for genocide*. Oxford, USA: Oxford Scholarship Online.

Lima, C. C. (2016). Os aliados ocultos de Hitler. *Superinteressante*. 31/10/2016. Disponível em: https://super.abril.com.br/historia/os-aliados-ocultos-de-hitler/ Acesso em: 23 set.2019.

Medeiros, C. R. O., & Alcadipani, R. (2018). Crimes corporativos: Uma leitura pós-colonial das noções formuladas na perspectiva funcionalista. *XVII Encontro Nacional de Pós-Graduação em Administração*, Curitiba, PR.

Mokhiber, R. (1995). *Crimes corporativos*. São Paulo, SP: Scritta, Página Aberta.

Mokhiber, R. (2007, January/February). 20 things about corporate crime. *Multinactional Monitor*. Recuperado de http://www.multinationalmonitor.org/mm2007/012007/mokhiber.html

Naimark, N. (2015, July 16). Genocide isn't history: It's part of the long-term human experience. *The Conversation*. Recuperado de http://theconversation.com/genocide-isnt-history-its-part-of-the-long-term-human-experience-40013

Pearce, F., & Tombs, S. (1999). *Toxic capitalism: Corporate crime and the chemical industry*. Toronto, Canada: Canadian Scholars' Press.

Sargent, N. C. (1990). Law, ideology and social change: An analysis of the role of law in the construction of corporate crime. *The Journal of Human Justice*, *1*(2), 97-116.

Schabas, W. (2009). *Genocide in international law: The crime of crimes*. Galway, Ireland: Cambridge.

Snider, L. (1990). Cooperative models and corporate crime: Panacea or cop-out. *Crime & Delinquency*, *36*, 373-390. doi: 10.1177/0011128790036003005

Stel, N. (2014, September 5). Business in genocide: Understanding the how and why of corporate complicity in genocides. *MsM conference Global Business, Emerging Markets and Human Rights*, Maastricht, Germany.

Stokes, P., & Gabriel, Y. (2010). Engaging with genocide: The challenge for organization and management studies. *Organization*, *17*(4), 461-480. doi: 10.1177/1350508409353198

Sutherland, E. H. (1940). White-collar criminality. *American Sociological Review*, *5*(1), 1-12. doi: 10.2307/2083937

Tombs, S., & Whyte. D. (2015). *The corporate criminal: Why corporations must be abolished*. London, UK: Routledge.

Unnever, J. D., Benson, M. L., & Cullen, F. T. (2008). Public support for getting tough on corporate crime: Racial and political divides. *Journal of Research in Crime and Delinquency*, *45*(2), 163-190. doi: 10.1177/0022427807313707

Whyte, D. (2018). The autonomous corporation: The acceptable mask of capitalism. *King's Law Journal*, *29*(1), 88-110. doi: 10.1080/09615768.2018.1475847

 CC BY