RAE Revista de Administração de Empresas®

PESQUISA E CONHECIMENTO
V. 59, N. 6,
Novembro–Dezembro 2019
fgv.br/rae

FGV EAESP

FÓRUM

A jornada acaba de começar
William Lekse

Além da tecnologia: Desafios gerenciais na era do *Big Data*
Eduardo de Rezende Francisco | José Luiz Kugler | Soong Moon Kang | Ricardo Silva | Peter Alexander Whigham

Capacidade de gestão da informação e implementação de estratégia de *Big Data*
Antonio Carlos Gastaud Maçada | Rafael Alfonso Brinkhues | José Carlos da Silva Freitas Junior

Intenção de adoção de *big data* na cadeia de suprimentos: Uma perspectiva brasileira
Maciel M. Queiroz | Susana Carla Farias Pereira

Medindo a acessibilidade: Uma perspectiva de *Big Data* sobre os tempos de espera do serviço da Uber
André Insardi | Rodolfo Oliveira Lorenzo

Fatores que afetam a adoção de análises de *Big Data* em empresas
Juan-Pedro Cabrera-Sánchez | Ángel F. Villarejo-Ramos

PERSPECTIVAS

***Big Data* e disrupções nos modelos de negócios**
Eric van Heck

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

PENSATA

Crimes corporativos: O espectro do genocídio ronda o mundo
Cintia Rodrigues de Oliveira

ISSN 0034-7590
0059 6
9 770034 759007

SUMÁRIO

EDITORIAL

372 PROPRIEDADE DOS DADOS E CIÊNCIA ABERTA
Data and open science
Propiedad de datos y ciencia abierta
Maria José Tonelli | Felipe Zambaldi

FÓRUM / *FORUM* / *FORO*

374 A JORNADA ACABA DE COMEÇAR
The journey has just begun
El viaje acaba de empezar
William Lekse

375 ALÉM DA TECNOLOGIA: DESAFIOS GERENCIAIS NA ERA DO BIG DATA
Beyond technology: Management challenges in the Big Data era
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 CAPACIDADE DE GESTÃO DA INFORMAÇÃO E IMPLEMENTAÇÃO DE ESTRATÉGIA DE BIG DATA
Information management capability and Big Data strategy implementation
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 INTENÇÃO DE ADOÇÃO DE BIG DATA NA CADEIA DE SUPRIMENTOS: UMA PERSPECTIVA BRASILEIRA
Intention to adopt big data in supply chain management: A Brazilian perspective
Intención de adopción de big data en la cadena de suministros: Una perspectiva brasileña
Maciel M. Queiroz | Susana Carla Farias Pereira

402 MEDINDO A ACESSIBILIDADE: UMA PERSPECTIVA DE BIG DATA SOBRE OS TEMPOS DE ESPERA DO SERVIÇO DA UBER
Measuring accessibility: A Big Data perspective on uber service waiting times
Medición de accesibilidad: Una perspectiva de Big Data sobre los tiempos de espera del servicio de la Uber
André Insardi | Rodolfo Oliveira Lorenzo

415 FATORES QUE AFETAM A ADOÇÃO DE ANÁLISES DE BIG DATA EM EMPRESAS
Factors affecting the adoption of Big Data analytics in companies
Factores que afectan a la adopción del análisis de Big Data em las empresas
Juan-Pedro Cabrera-Sánchez | Ángel F. Villarejo-Ramos

PERSPECTIVAS / *PERSPECTIVES*

430 BIG DATA E DISRUPÇÕES NOS MODELOS DE NEGÓCIOS
Big Data and disruptions in business models
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

PENSATA / *ESSAY* / *ENSAYO*

435 CRIMES CORPORATIVOS: O ESPECTRO DO GENOCÍDIO RONDA O MUNDO
Corporate crimes: The specter of genocide haunts the world
Crimen corporativos: El espectro del genocídio alrededor del mundo
Cintia Rodrigues de Oliveira

© RAE | São Paulo | **V. 59** | **n. 6** | **nov-dez 2019**
ISSN 0034-7590; eISSN 2178-938X

EDITORIAL

Versão original
DOI: http://dx.doi.org/10.1590/S0034-759020190601



Maria José Tonelli
Editora-chefe



Felipe Zamboldi
Editor-adjunto

PROPRIEDADE DOS DADOS E CIÊNCIA ABERTA

A prática de acesso aberto aos artigos, adotada no Brasil e em vários outros países, ainda enfrenta resistência por parte de muitas editoras comerciais no exterior (Packer & Santos, 2019a). Mais recentemente, algumas passaram a utilizar um modelo híbrido de acesso aberto e fechado, como um caminho alternativo e não tão polarizado. Também está em andamento a discussão sobre transparência no processo de revisão por pares: o SciELO recomenda "avanço gradual de transparência e abertura [...] com a abertura das identidades dos autores e pareceristas durante o processo de avaliação" (Packer & Santos, 2019b). Mas ainda mais polêmica é a política de acesso aberto dos dados das pesquisas dos artigos veiculados em publicações científicas. No *blog* SciELO em Perspectiva, encontram-se vários textos sobre essa tendência, questionada por diversos atores envolvidos no processo de produção e publicação de artigos científicos: autores, universidades, editores e *publishers*. Quem é o proprietário dos dados? Nassi-Calò (2019) mostra que as pesquisas sobre essa questão permanecem inconclusivas, e muitos atores desse processo podem ser os proprietários: financiadores da pesquisa, instituição do pesquisador, *publisher*, além, claro, da visão de que os dados são de propriedade dos autores. A autora argumenta que a ciência aberta "é uma demanda da sociedade, dos governos e financiadores. Esta prática traz inúmeras vantagens ao tornar a ciência mais transparente, reprodutível, confiável e verificável" (Nassi-Calò, 2019). Na perspectiva dos pesquisadores, entretanto, há inúmeras questões. Na pesquisa qualitativa, por exemplo, realizada por meio de entrevistas, quando se garante o anonimato dos entrevistados, registrado por meio de consentimento informado, como proceder? Nas Ciências Exatas e Biológicas, talvez essa questão não se coloque, mas é premente quando se trata de pesquisa nas Ciências Humanas, pois os informantes poderão ser identificados e o sigilo garantido dentro dos padrões éticos em pesquisa, violado. Isso sem considerar as questões de tempo e recursos dos pesquisadores, além da propriedade de dados secundários de terceiros que muitas vezes cedem o acesso exclusivamente para uma pesquisa específica. Outros aspectos que se fazem presentes para a transparência dos dados nessa era de ciência aberta, descrita como *e-science*, como apontam Targino e Garcia (2018), são a necessidade de ciberestrutura (bases tecnológicas que comportem os dados), a colaboração da sociedade, bem como o apoio do Estado. Mas, novamente, quais são os proprietários da infraestrutura tecnológica que guarda os dados? Packer e Santos (2019b) argumentam que ciência aberta é um movimento irreversível, e o 4.º Plano Brasileiro de Ação nesse tópico envolve alguns marcos claramente definidos para o futuro (Packer & Santos, 2019b), a partir das diretrizes da Global Open Fair. No Brasil, a área da saúde já trabalha com essas diretrizes, mas é necessário, argumentam os autores, que programas de pós-graduação invistam em programas de treinamento. Também a Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) trabalha com essa orientação em Projetos Temáticos. A valorização das pesquisas no futuro estará vinculada não só ao periódico em que o artigo foi publicado, mas também aos dados disponibilizados (Kiley & Markie, 2019). Supostamente, essa ação eliminaria problemas como plágio, reprodutibilidade da pesquisa, vieses, entre outros. Critérios que, sem dúvida, são válidos para as Ciências Exatas e Biológicas. Mas esses critérios são aplicáveis para as Ciências Humanas, que se ocupam, muitas vezes, de fenômenos únicos, não replicáveis? Se hoje até a neutralidade dos algoritmos é questionada, será que podemos mesmo fazer ciência sem vieses? Será que os dados são neutros?

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

Essa edição contempla justamente um fórum sobre *Big Data*, organizado por Eduardo de Rezende Francisco, José Luiz Kugler, Soong Moon Kang, Ricardo Silva e Peter Alexander Whigham. O fórum traz o primeiro artigo convidado, "A jornada acaba de começar", por William Lekse. Em seguida, a introdução ao fórum, "Além da tecnologia: Desafios gerenciais na era do *Big Data*", dos organizadores. E continua com os artigos: "Capacidade de gestão da informação e implementação de estratégia de *Big Data*", de Antonio Carlos Gastaud Maçada, Rafael Alfonso Brinkhues e José Carlos da Silva Freitas Junior; "Intenção de adoção de *big data* na cadeia de suprimentos: Uma perspectiva brasileira", de Maciel M. Queiroz e Susana Carla Farias Pereira; "Medindo a acessibilidade: Uma perspectiva de *Big Data* sobre os tempos de espera do serviço da Uber", de André Insardi e Rodolfo Oliveira Lorenzo, e "Fatores que afetam a adoção de análises de *Big Data* em empresas", de Juan-Pedro Cabrera-Sánchez e Ángel F. Villarejo-Ramos. A seção Perspectivas também traz o debate sobre o uso de *Big Data* nos negócios: "*Big Data* e disrupções nos modelos de negócios", por Eric Van Heck, e "*Plus ça change, plus c'est la même chose* [Quanto mais as coisas mudam, mais elas permanecem as mesmas] ", por Flávio Bartman. A pensata de Cintia Rodrigues de Oliveira, "Crimes corporativos: Um espectro ronda o mundo, o espectro do genocídio", nos lembra que más condutas, comportamentos antiéticos e irresponsabilidade social corporativa também permeiam o mundo dos negócios.

Boa leitura!

Maria José Tonelli[1] | ORCID: 0000-0002-6585-1493
Felipe Zambaldi[1] | ORCID: 0000-0002-5378-6444
[1] Fundação Getulio Vargas, Escola de Administração de Empresas de São Paulo, São Paulo, SP, Brasil

REFERÊNCIAS

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

FÓRUM

Artigo convidado
Versão traduzida
DOI: http://dx.doi.org/10.1590/S0034-759020190602

A VIAGEM ACABA DE COMEÇAR

Esta edição especial contribui para incorporar pesquisas que utilizam *Big Data* – principalmente as disciplinas de sistemas da informação e cadeia de suprimentos – na pesquisa acadêmica convencional. Também estende a contribuição do *Big Data*, permitindo que pesquisas baseadas em análise preditiva proponham teorias. O *Big Data* traz muitas características para a pesquisa acadêmica que, se devidamente compreendidas, podem mudar a abordagem da maioria das pesquisas para a abordagem acadêmica clássica, que se concentra na proposição e teste de teorias. A abordagem acadêmica para a realização de pesquisas teóricas busca explicar fenômenos por meio da aplicação de enquadramentos teóricos, os quais são oriundos de diferentes disciplinas como microeconomia, pesquisa operacional, teoria organizacional, psicologia e sociologia.

A capacidade de analisar, explicar e prever o comportamento do consumidor é de particular interesse para pesquisadores acadêmicos e profissionais. A maioria das vendas do varejo ainda ocorre nas lojas, e os consumidores que compram *on-line* também visitam lojas antes ou depois de uma venda. Atualmente, a maioria das pessoas que compram em lojas usa dispositivos móveis para pesquisar sobre produtos, se comunicar com familiares e amigos e visitar *sites*, os quais fornecem dados – geralmente, *Big Data* – com o objetivo de facilitar a experiência de compra (Fildes & Kolassa, 2018). Assim, não só as transações ou a falta delas, mas também todos os aspectos tecnológicos da experiência de compra, agora podem ser amplamente modelados. Grande parte desses dados está começando a ser disponibilizada para pesquisadores acadêmicos. Portanto, a tecnologia, principalmente a coleta de dados, processamento e disseminação de *Big Data*, está trazendo contribuições significativas para o mercado global.

Essas novas tecnologias e tendências em *Big Data* estão emergindo na análise local, regional e global do comportamento do consumidor e se estendendo ao longo das operações da cadeia de suprimentos. O *Big Data* está mudando as regras dos negócios – desde a concepção e prototipagem até a produção e distribuição de produtos e serviços. Agora, os acadêmicos têm à sua disposição o que já exigiam há muito tempo: fontes com quantidades enormes de dados. Durante décadas, as investigações sobre comportamento do consumidor publicadas em revistas foram limitadas a conjuntos de dados gerados pelos próprios pesquisadores. Os grandes conjuntos de dados de *Big Data* oferecem aos pesquisadores acadêmicos e profissionais os meios para se tornarem mais conscientes de atualizações relevantes em tempo real. Os pesquisadores não precisam mais gastar tempo desenvolvendo um plano de pesquisa – o que incluiria especificar enquadramentos, desenvolver modelos e buscar permissões – para conduzir investigações com pequenas amostras. Hoje, eles podem investigar vários modelos e enquadramentos de teorias integradas de várias disciplinas. Os meios para testar teorias explicativas e preditivas (incluindo teoria fundamentada) agora estão disponíveis para pesquisadores do mundo tudo. Além disso, todas as investigações podem ser rapidamente replicadas e verificadas, pois os dados estão disponíveis para todos os acadêmicos, permitindo um ambiente de pesquisa global mais produtivo e criativo (Johnson, Gray, & Sarker, 2019).

Esta edição especial explora diferentes métodos para enfrentar desafios analíticos relevantes. Essa emocionante jornada acaba de começar, e certamente levará a caminhos de pesquisa interessantes.

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, Estados Unidos da América

REFERÊNCIAS

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

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

FÓRUM

Artigo convidado
Versão traduzida
DOI: http://dx.doi.org/10.1590/S0034-759020190603

ALÉM DA TECNOLOGIA: DESAFIOS GERENCIAIS NA ERA DO *BIG DATA*

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, Brasil

[2] University College London, School of Management, Londres, Reino Unido

[3] University College London, Department of Statistical Science, Londres, Reino Unido

[4] University of Otago, Department of Information Science, Dunedin, Otago, Nova Zelândia

INTRODUÇÃO

A capacidade das organizações de produzir, coletar, gerenciar, analisar e transformar dados aumentou rapidamente na última década (Delen & Zolbanin, 2018). Isso gerou novos desafios significativos em relação a como os dados podem ser aproveitados para melhorar as decisões de negócios e como esse novo cenário altera os processos e as operações de negócios (Vidgen, Shaw, & Grant, 2017). A adoção generalizada de métodos analíticos avançados (por exemplo, aprendizado de máquina) tem atraído bastante interesse (Gupta, Deokar, Iyer, Sharda, & Schrader, 2018; Vassakis, Petrakis, & Kopanakis, 2018), principalmente porque o armazenamento de dados e os métodos necessários podem ser acessados remotamente por meio de interfaces baseadas na *web*, como serviços em nuvem. Isso gerou uma crença crescente de que as empresas devem envolver-se ativamente com essa tecnologia para se manterem competitivas. No entanto, esse cenário de corrida da Rainha Vermelha (que pressupõe um desenvolvimento contínuo por parte das empresas) tem um custo, pois a coleta, a curadoria e o gerenciamento de grandes conjuntos de dados requerem experiência e uma equipe dedicada, o que, muitas vezes, consome recursos que não contribuem para as principais atividades do negócio. É preciso considerar também que cientistas de dados e engenheiros de dados, entre outros, cada vez mais exercem um papel relevante dentro das organizações (Davenport & Patil, 2012). Cargos como *Chief Data Officer* (CDO) e *Chief Analytics Officer* (CAO) agora são comuns na maioria das organizações.

Há também a questão da preparação de dados. O mantra de que 80% do esforço é concentrado no gerenciamento de dados ainda é amplamente válido. Além disso, a utilização e interpretação adequadas de modelos preditivos requerem conhecimentos especializados que envolvem tanto a compreensão do negócio subjacente como os pressupostos e limitações de cada modelo. Encontrar pessoas adequadas com habilidades tanto do ponto de vista empresarial quanto do ponto de vista tecnológico pode ser difícil. As empresas, muitas vezes, têm dificuldade de avaliar a relação custo-benefício do *Big Data*, o que pode levar a falhas na forma como o *software* é desenvolvido e vinculado ao modelo de negócio (Loebbecke & Picot, 2015). Existem muitos exemplos de organizações, principalmente organizações governamentais e financiadas com recursos públicos, em que recursos escassos são desperdiçados todo ano em projetos analíticos que falham devido a um mal-entendido sobre como os dados devem ser usados, os tipos de dados que são coletados e as questões que o modelo se propõe a resolver. Há também questões relevantes relacionadas a ciclos de vida de desenvolvimento lento na arena analítica. Uma vez que a tecnologia está evoluindo em um ritmo

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

acelerado, um projeto que leva um tempo significativo para ser concluído pode resultar em uma solução cara em comparação a uma solução que usa tecnologia mais moderna. Saber quando desenvolver e quando esperar é também um desafio fundamental para os mecanismos atuais de governança analítica.

Há uma necessidade de entender como as organizações devem transformar seus modelos de negócios quando confrontadas com esse mundo cada vez mais rico em termos de dados, e como elas podem garantir conformidade com práticas corretas não apenas do ponto de vista tecnológico, mas também dos pontos de vista gerencial, ético e social. As primeiras discussões sobre o tema do *Big Data* foram, muitas vezes, enquadradas dentro da perspectiva do V (volume, velocidade, variedade, valor, veracidade, variabilidade, visualização) (Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015), e, embora esses conceitos ainda permaneçam relevantes, é cada vez mais reconhecido que os dados não são um conceito desconectado. Isso levou ao entendimento do gerenciamento de dados a partir de uma perspectiva ecossistêmica (Demchenko, Laat, & Membrey, 2014). Em termos gerais, um ecossistema natural opera em uma variedade de escalas espaciais e temporais, abrangendo desde o indivíduo dentro de uma espécie até a comunidade de espécies, teias alimentares e o meio ambiente, tudo dentro do contexto de fatores exógenos (como clima e competição) e endógenos (tais como requisitos de nutrientes). Os dados também podem ser entendidos dentro desse enquadramento mais amplo (Gupta et al., 2018), devendo, portanto, ser utilizados e modelados como parte de um sistema mais amplo e dinâmico, em vez de um conceito separado e desconectado. Isso inclui a origem da coleta dos dados, outros dispositivos digitais e sensores, provedores de tecnologia e comunidades mais amplas envolvidas na criação de dados, elaboração de políticas, e assim por diante.

Quais são as tendências atuais na análise de *Big Data*? Existem dois caminhos principais que vale a pena mencionar em relação à tomada de decisão empresarial: infraestruturas de dados integradas (IDIs) e a internet das coisas (IoT) (Ahmed et al., 2017). A ideia principal das IDIs é que a ligação ou associação de dados pode proporcionar oportunidades extras para examinar a estrutura e as relações entre todos esses conjuntos de dados. Muitas organizações governamentais coletaram dados como justiça, saúde, educação, renda, serviços sociais, comunidade e estatísticas populacionais (por exemplo, coletas regulares de censos) por meio de órgãos separados. No entanto, até recentemente, a maioria desses dados não podia ser vinculada de modo útil, e era difícil organizá-los em formato comum ou obter acesso a esses tipos de dados. As IDIs permitem que esses tipos de dados sejam usados em conjunto, permitindo a emergência de uma visão ecossistêmica da sociedade contendo microdados centrados na pessoa e que podem estar relacionados a dados escalados e agregados. Compreender como os indivíduos interagem, como tomam decisões e quais resultados geram na sociedade (Newell & Marabelli, 2015) permite uma maior compreensão de por que as pessoas se comportam de certa forma em diferentes circunstâncias. Isso significa que as empresas devem compreender a forma como as estruturas sociais, desde o indivíduo à perspectiva de comunidade, estão operando e, portanto, as oportunidades de negócios que podem ser aproveitadas a partir da perspectiva individual. As IDIs permitem a abordagem de questões de áreas tão diversas como produção agrícola, saúde mental, desenvolvimento da educação, mercado de trabalho, imigração, turismo, disparidades salariais e desigualdade de gênero. No entanto, usar uma IDI gera complicações como questões de segurança, uma vez que o acesso é, muitas vezes, limitado ou altamente controlado, e o acesso para fins comerciais pode também ser limitado, a menos que haja uma associação direta com uma organização de pesquisa, como uma universidade. No entanto, o atual entendimento no desenvolvimento de IDIs é de que as empresas, em última análise, se beneficiarão dessas fontes de dados conectadas, quer seja para fins próprios, quer como fornecedoras de ferramentas e métodos voltados para integração e utilização dessas fontes de dados.

A IoT (Ahmed et al., 2017) continua a ser impulsionada pela demanda do consumidor com a promessa de prover uma melhor personalização dos serviços e controle sobre muitos processos de tomada de decisão individuais. Uma eventual implementação de redes sem fio 5G rápidas e o aumento da conectividade entre todos os dispositivos (vale lembrar quando o *smartphone* foi introduzido, mas agora é apenas um telefone) gerariam oportunidades de negócios em relação ao modo como essas conexões são usadas, o que elas representam de uma perspectiva individual, e quais novos produtos e serviços podem ser criados em torno desse ecossistema. O acesso a esses dados também propiciará o surgimento de novas abordagens para entender o comportamento individual, como a demanda do consumidor é criada (Erevelles, Fukawa, & Swayne, 2016) e métodos para otimizar como os indivíduos interagem com os sistemas. Os dispositivos eletrônicos inteligentes tornam o processamento local possível; a noção de computação de ponta e o pré-processamento de dados para filtrar a forma como a informação é usada se tornarão aspectos fundamentais do desenvolvimento da IoT. As oportunidades de negócios existem tanto do ponto de vista de *hardware* quanto do ponto de vista

de *software*, abrangendo desde quais tipos de dispositivos serão usados até a forma como eles interagirão como um sistema. A discussão sobre o mundo dos sensores e como isso mudará nossa perspectiva em termos de oportunidades de negócios apenas começou.

Esse influxo enorme e sem precedentes de dados oferece uma lista infinita de oportunidades. No entanto, para alavancar tais oportunidades, precisamos desenvolver modelos significativos; para dar sentido à complexidade que caracteriza nossos desafios econômicos, políticos e sociais, precisamos desenvolver modelos bem-articulados e que façam sentido, ao mesmo tempo que revelem como os processos causais se sobrepõem e interagem (Page, 2018).

Do ponto de vista gerencial, a missão é reconhecer quais processos de negócios podem se beneficiar de que tipo de modelos, como os dados podem ser organizados e utilizados e como os resultados analíticos podem ser incorporados ao quadro de tomada de decisões.

ARTIGOS ACEITOS

Nesta edição especial, focamos o mais básico desses desafios, a saber, a decisão de implementar tecnologias de *Big Data*. Em "Fatores que afetam a adoção da análise de *big data* nas empresas", Cabrera-Sánchez e Ramos (2019) examinam as barreiras para a implementação de técnicas de *Big Data* com base em pesquisas *on-line* com gestores de diferentes áreas, como *Marketing*, Finanças e Recursos Humanos. Eles descobriram que empresas com pouca ou nenhuma experiência com *Big Data* são mais propensas à influência social, têm maiores expectativas em relação à nova tecnologia e têm maior resistência para adotar a nova tecnologia, enquanto empresas com mais experiência estão mais interessadas em facilidade de acesso e obtenção de suporte necessário para uso da tecnologia, tendo expectativas mais baixas em relação a seu desempenho.

Dirigindo especial atenção a experiências no Brasil, Queiroz e Farias (2019) utilizam um marco semelhante ao empregado por Cabrera-Sánchez e Ramos (2019) em seu artigo "Intenção de adotar *Big Data* na gestão da cadeia de suprimentos: Uma perspectiva brasileira", nomeadamente a teoria unificada da aceitação e uso da tecnologia (Utaut), para analisar especificamente a intenção de adotar técnicas de *Big Data* entre profissionais brasileiros que trabalham na gestão da cadeia de suprimentos e que tinham alguma experiência com a tecnologia. Para esses profissionais, a adoção da tecnologia de *Big Data* depende diretamente da infraestrutura de TI, que inclui acesso à internet de alta velocidade e a integração com outros sistemas.

Em seu artigo, "Capacidade de gerenciamento de informações e implementação de estratégias de *big data*", Maçada, Brinkhues e Freitas (2019) investigam como as expectativas de uma organização em relação aos benefícios e custos de *Big Data* são influenciadas por sua capacidade de acessar dados e informações em seu ambiente, de processá-los e de atender às necessidades do mercado com base neles, ou "information management capability" (IMC) – capacidade de gerir informações. Eles demonstram que a IMC está positivamente relacionada às expectativas de valor e negativamente relacionada às expectativas de custo, o que, por sua vez, afeta negativamente a intenção de adquirir recursos e capacidades para implementar *Big Data*.

Finalmente, em um estudo que faz uso de Big Data, Insardi e Lorenzo (2019), em "Medindo acessibilidade: Tempos de espera do serviço Uber sob a perspectiva do *Big Data*", utilizaram algumas técnicas de *Big Data* para estudar o acesso à mobilidade em um ambiente urbano de grande porte usando tempos de espera estimados de todos os produtos Uber na cidade de São Paulo. A principal constatação deles foi que os tempos de espera estimados estão fortemente relacionados às variáveis socioeconômicas dos distritos da cidade. Por exemplo, os autores encontraram uma forte relação entre os tempos de espera e a proporção de população não branca.

REFERÊNCIAS

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). Fatores que afetam a adoção de análises de *Big Data* em empresas. *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

 CC BY

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).Medindo a acessibilidade: Uma perspectiva de Big Data sobre os tempos de espera do serviço da Uber. *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). Capacidade de gestão da informação e implementação de estratégia de Big Data. *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). Intenção de adoção de big data na cadeia de suprimentos: Uma perspectiva brasileira. *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

FÓRUM

Submetido 01.10.2018. Aprovado 19.07.2019
Avaliado pelo sistema double blind review. Editores Científicos Convidados: Eduardo de Rezende Francisco, José Luiz Kugler, Soong Moon Kang, Ricardo Silva e Peter Alexander Whigham
Versão traduzida
DOI: http://dx.doi.org/10.1590/S0034-759020190604

CAPACIDADE DE GESTÃO DA INFORMAÇÃO E IMPLEMENTAÇÃO DE ESTRATÉGIA DE *BIG DATA*

Information management capability and Big Data strategy implementation

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

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 gestão da 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 a CGI como um agente de imperfeição no mercado de fatores estratégicos de *Big Data*. As hipóteses foram testadas a partir de uma Survey com 101 respondentes e analisadas com a utilização de SEM-PLS. Os resultados indicam uma influência CGI 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.

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.*

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.*

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, Brasil

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

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

 CC BY

INTRODUÇÃO

"*Big Data* é possivelmente a mais significativa ruptura de 'tecnologia' nos ecossistemas empresariais e acadêmicos desde a ascensão meteórica da Internet e da economia digital" (Agarwal & Dhar, 2014, p. 443). Diversas formas de dados que não geram valor não contribuem para as organizações. O valor dos dados está, portanto, impulsionando o crescente interesse em *Big Data* (Chiang, Grover, Liang, & Zhang, 2018). Pesquisadores e fornecedores de tecnologia reconhecem os benefícios da adoção da análise de *Big Data* em práticas de negócios (Wang, Kung, Wang, & Cegielski, 2018). As empresas estão cada vez mais interessadas em desenvolver estratégias de *Big Data* (Tabesh, Mousavindin, & Hasani, 2019). O percentual de empresas que já investem ou planejam investir em *Big Data* cresceu de 64% em 2013 (Gartner, 2014) para 73% em 2018 (Davenport & Bean, 2018). "As organizações estão atualmente buscando adotar a tecnologia de *Big Data*, mas estão incertas dos benefícios que ela pode trazer para a organização, além de se preocuparem com os custos de implementação" (Lakoju & Serrano, 2017, p. 1). O volume de investimentos está crescendo a um ritmo ainda maior. O mercado de tecnologia e serviços de *Big Data* crescerá a uma taxa de crescimento anual composta (TCAC) de 11,9%, atingindo 260 bilhões de dólares até 2022 (International Data Corporation [IDC], 2018). Os impactos organizacionais esperados são muitos, e incluem redução de custos, aumento nos *insights* de negócios, revelações de informações estratégicas e melhor tomada de decisão (Kwon, Lee, & Shin, 2014). No entanto, o valor esperado desses benefícios e os custos envolvidos na aquisição e desenvolvimento dessas soluções não são os mesmos para cada empresa, o que gera imperfeições competitivas no mercado de recursos estratégicos.

De acordo com a teoria do mercado de fatores estratégicos (MFE), as empresas precisam estar consistentemente mais informadas do que outras empresas que visam implementar a mesma estratégia para obter desempenho superior (Barney, 1986), este autor afirma que a análise das capacidades da empresa pode ajudar mais na criação dessas circunstâncias que o ambiente competitivo. Argumentamos que a capacidade de gestão da informação (CGI) pode trazer o único *insight* necessário para a implementação de estratégias bem-sucedidas de *Big Data*. Definimos CGI como a capacidade da empresa de acessar dados e informações de ambientes internos e externos, mapear e distribuir dados para processamento, e permitir que se ajuste para atender as necessidades e direções do mercado. A literatura indica que a CGI influencia positivamente e diretamente o desempenho de uma empresa (Carmichael, Palacios-Marques, & Gil-Pichuan, 2011) ou é mediada por outras capacidades organizacionais (Mithas, Ramasubbu, & Sambamurthy, 2011). Não há evidências de que a CGI atual de uma empresa possa acomodar o crescimento acentuado do fluxo de dados não estruturados (White, 2012).

No entanto, a CGI pode ter um papel relevante nas expectativas e intenção de implementar uma estratégia para lidar com *Big Data*. Muitos adeptos de *Big Data* estão buscando tais oportunidades devido ao fácil acesso a capacidades computacionais e *software* analíticos (Agarwal & Dhar, 2014). Por outro lado, 43% dos diretores referem-se aos défices orçamentários como a principal barreira atrapalhando ações que tirem proveito desse contexto (McKendrick, 2013). Isso indica simetria na expectativa de custo dos recursos para a implementação de uma estratégia de *Big Data*. Do ponto de vista acadêmico, muitos estudos investigam esse fenômeno, especialmente da área de Sistemas da Informação (SI), os quais analisam a criação de valor a partir desses dados (por exemplo, Brown, Chui, & Manyika, 2011; Davenport, Barth, & Bean, 2012; Johnson, 2012; Lakoju & Serrano, 2017; McAfee & Brynjolfsson, 2012).

No entanto, poucos trabalhos focam a relação entre CGI e *Big Data* para a obtenção desse valor (Brinkhues, Maçada, & Casalinho, 2014; Mohanty, Jagadeesh, & Srivatsa, 2013). "A literatura atual sobre a realização do valor de *Big Data* é caracterizada por um número limitado de estudos empíricos e alguns que trazem ideias antigas sob novos formatos" (Günther, Rezazade Mehrizi, Huysman, & Feldberg, 2017). Este estudo tem como objetivo determinar como a variação no nível de CGI entre as empresas cria imperfeições competitivas no mercado de recursos para a implementação de estratégias de *Big Data*. Para cobrir essa lacuna de pesquisa, propomos uma escala para medir a CGI e desenvolver conceitualmente um modelo de pesquisa para avaliar empiricamente a relação entre a CGI e a implementação da estratégia de *Big Data*. Esse modelo, baseado na teoria do MFE, investiga especificamente a influência da CGI sobre o valor e as expectativas de custo dos recursos necessários para essa implementação, e, com base na teoria do custo de transação, o efeito dessas expectativas sobre a intenção de adquirir ou desenvolver esses recursos. Construímos a escala seguindo a literatura e coletamos dados de executivos via classificação de cartões. O modelo de pesquisa foi testado por meio de uma *Survey* com 101 diretores, e os dados foram analisados utilizando o SEM-PLS.

O artigo prossegue como detalhado a seguir. A próxima seção desenvolve as hipóteses e apresenta o modelo de pesquisa. A seção seguinte detalha os procedimentos para construir a escala de CGI e para a coleta de dados. Posteriormente, apresentamos e discutimos os resultados e, finalmente, apresentamos nossas conclusões e implicações para a pesquisa e prática gerencial.

O CGI E O MFE

"Mercados de Fatores Estratégicos (MFE) são mercados onde os recursos necessários para a implementação de uma estratégia são adquiridos" (Barney, 1986, p. 1231); assim, as empresas só podem extrair desempenho superior quando o MFE for imperfeito devido às diferenças na expectativa do valor futuro desses recursos estratégicos. Em outras palavras, as empresas devem poder extrair um maior valor dos recursos necessários para a sua implementação estratégica em vez de arcar com custos de aquisição significativamente inferiores ao seu valor econômico. "O objetivo dos programas de *Big Data* deve ser fornecer valor suficiente para justificar que continuem sendo utilizados na exploração de novas capacidades e *insights*" (Mithas, Lee, Earley, Murugesan, & Djavanshir, 2013, p. 18). Para obter essa vantagem, as empresas têm de estar mais bem-informadas do que as outras empresas que atuam no mesmo MFE (Barney, 1986). A CGI pode servir como um propulsor dessa vantagem.

Mithas et al. (2011) propõem o construto CGI para desenvolver um modelo conceitual vinculando-o a três outras capacidades organizacionais (gestão de clientes, gestão de processos e gestão de desempenho). Seus resultados mostram que essas capacidades de gestão mediam a influência positiva da CGI no desempenho da empresa. O conceito de CGI de Mithas et al. (2011) consiste em três habilidades: fornecer dados e informações aos usuários a partir de níveis adequados de precisão, pontualidade, confiabilidade, segurança e confidencialidade; fornecer conectividade e acesso universal com escopo e escala adequados; e adaptar a infraestrutura às necessidades emergentes e direções do mercado. Carmichael et al. (2011) definem a CGI como um construto de segunda ordem composto pela compilação e produção de informações; acesso à informação; e identificação de requisitos de distribuição de informações. Outro autor, Phadtare (2011), propõe que a CGI está ligada a cinco fatores: aquisição e retenção, processamento e síntese, recuperação e uso, transmissão e divulgação, e sistema de apoio e integração.

Com base nos três trabalhos mencionados acima (Carmichael et al., 2011; Mithas et al., 2011; Phadtare, 2011), identificamos cinco dimensões da CGI (acesso, distribuição, pessoas, arquitetura e infraestrutura). Em seguida, como explicaremos detalhadamente nas próximas seções, realizamos uma análise de classificação de cartões com executivos, que apontou para uma escala composta por 10 itens dessas dimensões. A partir dessa análise, formulamos uma definição de CGI que, aplicada neste estudo, corresponde ao conjunto de habilidades da empresa que articulam a infraestrutura da informação, a arquitetura da informação e o acesso à informação, possibilitando o ajuste organizacional em resposta às mudanças impostas por ambientes externos. Assim, esperamos que as organizações com CGI mais desenvolvida sejam mais precisas em suas expectativas de valor e possam tirar proveito da assimetria da informação no MFE, de onde derivam imperfeições competitivas no MFE.

Além disso, esperamos que as empresas que desenvolveram CGI com uma qualidade superior durante uma das eras anteriores de GI — Suporte a Decisão, Suporte Executivo, Processamento Analítico *On-line* e Inteligência e Análise de Negócios (Davenport, 2014) — tenham uma maior expectativa de valor da próxima fronteira do *Big Data*. Prevemos esse resultado porque o desenvolvimento da CGI em nível elevado impacta positivamente o desempenho organizacional (Carmichael et al., 2011; Mithas et al., 2011), o que favorece um efeito polarizador das percepções entre o passado e o presente (Vasconcelos, Mascarenhas, & Vasconcelos, 2006). A estratégia de *Big Data* é um conjunto de soluções baseadas em avanços recentes na análise de *Big Data*. As organizações buscam incorporar essas soluções em seus próprios processos de tomada de decisão com sucesso (Tabesh et al., 2019). Assim, essas empresas têm uma maior expectativa de valor das estratégias de *Big Data* com base em suas experiências anteriores positivas com investimentos de GI. Por outro lado, as empresas que não atingiram o mesmo nível de CGI podem não ter tido o mesmo sucesso nos seus empreendimentos em GI, e essa experiência negativa pode resultar numa maior expectativa de custo para adotar esse tipo de estratégia.

> H1: As empresas com CGI mais elevada têm uma expectativa de custo menor para implementar uma estratégia de *Big Data*.
>
> H2: As empresas com CGI mais elevada têm maiores expectativas de extração de valor da implementação de uma estratégia de *Big Data*.

Expectativa de valor assimétrico e intenção de adquirir/desenvolver capacidades de estratégia de *Big Data*

Estudos anteriores também demonstram o efeito positivo da utilização de dados para a aquisição de soluções de *Big Data* (Kwon et al., 2014). No entanto, as empresas também podem desenvolver os recursos e capacidades para implementar uma estratégia de *Big Data* internamente.

As organizações existem para realizar transações internas de modo mais eficiente do que seria realizá-las no mercado (Coase, 1937). Por conseguinte, as empresas que não organizam os seus recursos para atingir os seus objetivos de maneira mais eficiente do que o mercado perdem a sua razão de existir. Assim, a busca dos recursos necessários para implementar uma

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

estratégia de *Big Data* pode tomar dois caminhos: desenvolvê-los internamente ou adquiri-los no mercado. As organizações podem desenvolver internamente as capacidades necessárias para essa implementação se forem eficientes na reorganização dos recursos envolvidos. No entanto, se o custo para adquirir esses fundos no mercado for menor do que o valor para produzi-los internamente, as empresas tendem a adquiri-los.

Os custos das transações são consequência da distribuição assimétrica e incompleta da informação entre as organizações envolvidas na troca (Cordella, 2006). O surgimento de vários fornecedores com soluções para gerenciar *Big Data* provoca incerteza sobre o valor que as empresas podem extrair desses recursos. Assim, a decisão de adquirir ou desenvolver os fatores necessários para implementar uma estratégia de *Big Data* também é afetada pelas diferenças nas expectativas assimétricas de valor que a empresa pode extrair desse investimento. Acreditamos que diferentes níveis de expectativas influenciam positivamente ambas as decisões, seja a de adquirir ou a de desenvolver internamente os recursos para extrair valor de *Big Data*.

> H3a: Empresas com maiores expectativas de extração de valor das estratégias de *Big Data* têm maior intenção de adquirir essas soluções.
>
> H3b: Empresas com maiores expectativas de extração de valor de *Big Data* têm maior intenção de desenvolver essas soluções internamente.

Expectativa de custos assimétricos e intenção de adquirir/desenvolver capacidades de estratégia de *Big Data*

Recursos como milhões de instruções por segundo (MIPS) e *terabytes* de armazenamento para dados estruturados são menos dispendiosos quando operados por meio de tecnologias de *Big Data* do que por meio de tecnologias tradicionais (Davenport, 2014). No entanto, os custos de outros recursos menos tangíveis podem ser mais difíceis de prever.

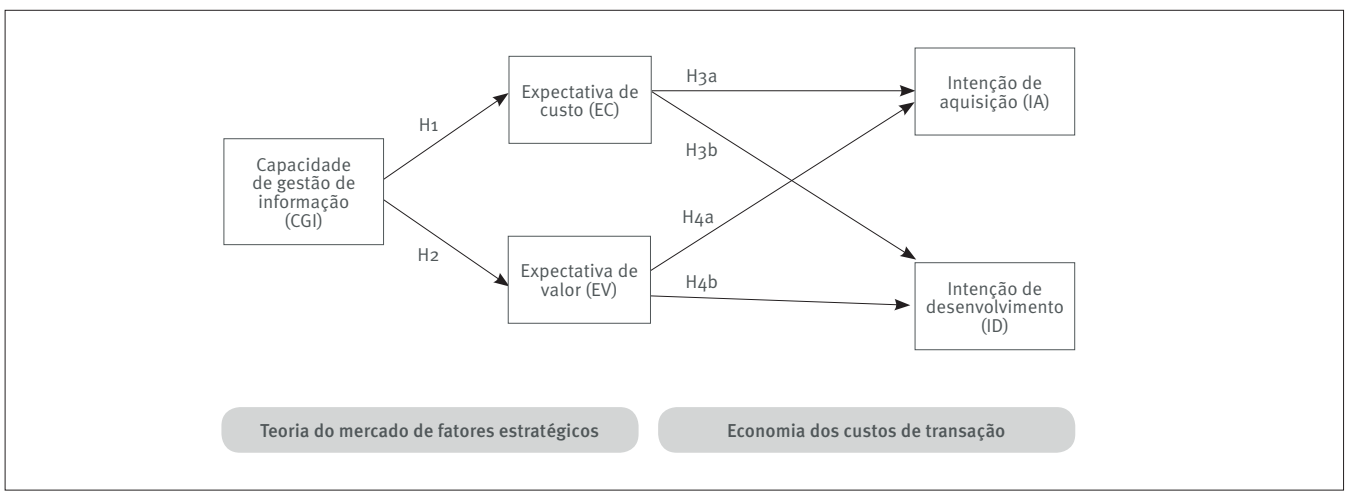
Por exemplo, os custos de transação frequentemente aumentam ao adotar uma solução de SI. No entanto, quando os custos associados à adoção não excedem os custos externos que afetam a adoção, as empresas podem reduzi-los (Cordella, 2006).

Assim como acreditamos que empresas com CGI mais bem desenvolvida tenham uma menor expectativa dos custos necessários para adotar uma estratégia de *Big Data*, também é provável que essa previsão de custos reduzidos aumente a predisposição para a implementação. Além disso, com uma expectativa de custo mais precisa, as empresas com um nível elevado de CGI podem criar uma estratégia adequada a seus orçamentos. Acreditamos que o efeito oposto também é válido: as empresas com CGI menos desenvolvidas tendem a ter previsões de custos menos exatas e, portanto, maior incerteza ao decidir se compram ou desenvolvem recursos para implementar uma estratégia de *Big Data*.

> H4a: Empresas com maiores expectativas dos custos para implementar estratégias de *Big Data* têm menor intenção de adquirir essas soluções.
>
> H4b: Empresas com maiores expectativas do custo para implementar estratégias de *Big Data* têm menor intenção de desenvolver essas soluções internamente.

Considerando as quatro hipóteses desenvolvidas acima, construímos o modelo de pesquisa. Uma ilustração disso pode ser vista na Figura 1.

Figura 1. **Modelo de pesquisa**



METODOLOGIA DE PESQUISA

As hipóteses foram testadas utilizando uma modelagem de equações estruturais de mínimos quadrados parciais (PLS-SEM) com base nos dados da *Survey*. PLS-SEM é frequentemente recomendado para pesquisa em gestão, porque os dados nesse campo muitas vezes não aderem a uma distribuição normal multivariada, ao mesmo tempo que os modelos são complexos e podem ser informativos. Também é recomendado para amostras menores e modelos com menos suporte prévio na literatura (Hair, Hult, Ringle, & Sarstedt, 2016; Ringle, Silva, & Bido, 2014). Diante das variáveis envolvidas e da natureza desta pesquisa, consideramos o uso dessa técnica estatística apropriada para testar empiricamente as hipóteses do modelo conceitual.

No entanto, realizamos uma etapa preliminar que consistiu na aplicação de um questionário e análise de *Card Sorting* para propor uma escala para mensurar CGI. Descrevemos essa etapa na próxima seção, e em seguida descrevemos os passos e detalhes sobre a amostra, coleta de dados e validação.

Classificação de cartões para criar uma escala de CGI

Adaptamos uma escala para medir a CGI na fase quantitativa usando uma *Survey*. Essa escala foi baseada em instrumentos de pesquisa existentes (Carmichael et al., 2011; Mithas et al., 2011). A necessidade de construir uma escala de CGI que pudesse lidar com esse novo ambiente de dados não influenciou as demais variáveis, que já tinham escalas testadas.

Para a escala, aplicamos a ferramenta Optimal Workshop para realizar uma classificação de cartões com 10 executivos de TI. Cada participante *on-line* levou uma média de sete minutos para concluir. Com base nos resultados da classificação de cartões, reduzimos a escala de 20 itens em cinco dimensões (pessoas, distribuição, acesso, infraestrutura e arquitetura de informação) para 10 itens, analisando uma matriz na qual usamos o corte acima de 60% de similaridade. Para avaliar as dimensões, utilizou-se uma análise de dendrograma para o melhor método de mesclagem, que frequentemente supera o método de concordância real quando uma pesquisa tem menos participantes. Ele faz suposições sobre *clusters* mais maciços com base em relacionamentos de pares individuais (Optimal Workshop, 2017). As pontuações do corte representam 40% dos participantes que concordam com partes desse agrupamento. Cinco dimensões emergiram do grupo de escalas-itens avaliados pelos executivos, que, por sua vez, foram selecionados a partir da literatura existente. Esse grupo foi coletado por meio da análise de classificação de cartões e nomeado com base nos itens coletados (pessoas, distribuição, acesso, infraestrutura e arquitetura da informação), de acordo com a análise dos autores dos resultados da etapa preliminar do estudo.

Assim, desenvolvemos a escala de CGI para este estudo. Desenvolvemos essa escala porque uma pesquisa aprofundada sobre esse construto (Mithas et al., 2011) foi validada a partir de uma adaptação utilizando dados secundários preexistentes e também para incorporar elementos abordados em outros trabalhos (Carmichael et al., 2011). As escalas para as demais variáveis do instrumento de pesquisa foram adaptadas da literatura e modificadas conforme necessário para este estudo. Todos os itens usaram uma escala Likert de sete pontos (1-Discordo fortemente; 7-Concordo fortemente). A análise estatística foi realizada utilizando o pacote de *software* SmartPLS versão 3.2.0.

Exemplo de quadro e coleta de dados

Coletamos dados por meio de uma pesquisa *on-line* criada usando a plataforma Google Forms. Os dados foram coletados em redes sociais, principalmente de grupos específicos de discussão sobre os temas abordados. Cerca de 29.282 pessoas viram os avisos, 208 pessoas clicaram neles, e recebemos 114 formulários preenchidos. A taxa de resposta foi de 59%. Entre estas, 13 foram eliminadas por meio de três questões de validação inseridas no questionário para auxiliar o controle da qualidade dos dados, o que nos deixou com uma amostra final de 101 formulários. Assim, a amostra excede o mínimo de 68 casos considerando um poder de 0,8 e um tamanho de efeito médio f^2 de 0,15 (Hair et al., 2016) com as variáveis tendo um número máximo de dois preditores. A amostra mínima foi calculada usando a ferramenta G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009).

Os entrevistados eram gerentes e executivos de TI ou de outras áreas relacionadas à implementação de estratégias de GI. A Tabela 1 resume os perfis das empresas respondentes, a partir do qual podemos concluir que a amostra é diversificada e levemente focada na indústria e tamanho, seja por meio do número de funcionários ou faturamento. As duas diferenças mais aparentes na variável de tamanho aparecem nas duas primeiras linhas. Na primeira linha, há uma porcentagem menor de empresas faturando até um milhão de dólares (16%), enquanto a porcentagem de empresas com até 50 funcionários é de 27%. Em contrapartida, a segunda linha apresenta um percentual maior de faturamento (23% de 1 a 6,7 milhões de dólares) e um número menor de empregados. Uma possível explicação

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

para essas diferenças podem estar no elevado número de postos de trabalho tecnológicos, os quais têm um alto potencial de rentabilidade, mesmo tendo menos trabalhadores. Houve diferenças significativas nos resultados relativos ao tamanho da indústria ou da empresa. Ao utilizar técnicas de mistura finita em PLS, não foram identificadas classes latentes que evidenciem a presença de grupos dentro de uma amostra.

Tabela 1. **Perfil das empresas respondentes**

Indústria	%	Número de Funcionários	%	Receita Anual	%
Tecnologia	24%	Até 50	27%	Até 1 milhão de dólares	16%
Manufatura	18%	51-100	13%	de 1 a 6,7 milhões de dólares	23%
Serviços financeiros	12%	101-500	11%	de 6,7 a 37,5 milhões de dólares	14%
Serviços profissionais	11%	501-1.000	16%	de 37,5 a 125 milhões de dólares	12%
Outros	35%	Mais de 1.000	33%	Mais de 125 milhões de dólares	36%

Nota: n=101

RESULTADOS

Primeiramente, apresentamos uma análise dos resultados em termos do modelo de medição, seguida de uma avaliação do modelo estrutural.

Avaliação do modelo de medição

O modelo de medição foi avaliado por meio de uma série de testes de confiabilidade, incluindo confiabilidade composta (CC), alfa de Cronbach, variância média extraída (VME) e validade discriminante (Hair et al., 2016; Ringle et al., 2014). Como mostra a Tabela 2, seguindo os critérios de Fornell e Larcker (Henseler, Ringle, & Sinkovics, 2009), o modelo converge, e o resultado é satisfatório, porque a VME de todas as variáveis está acima de 0,50.

Embora o indicador tradicional para avaliar a consistência interna seja o alfa de Cronbach, CC é o melhor indicador para PLS-PM, pois é o menos sensível ao número de itens em cada construto (Ringle et al., 2014). Na Tabela 2, também vemos que todas as variáveis apresentam ambos os indicadores (alfa de Cronbach e CC) acima de 0,7. Portanto, todas as variáveis são consideradas adequadas e satisfatórias (Hair et al., 2016). Ainda na Tabela 2, relatamos os critérios de Fornell e Larcker (1981) para verificar a qualidade discriminante com base nos valores de correlação entre as variáveis. Os resultados indicam que não há correlação entre variáveis distintas maiores que a raiz quadrada da VME de cada variável (destacadas em cinza na diagonal principal).

Como último critério para avaliar a qualidade do modelo de medida, a validade discriminante foi calculada utilizando uma análise de cargas cruzadas (Chin, 1998). Na Tabela 3, não encontramos indicadores com cargas fatoriais abaixo de sua variável em relação aos demais. Tendo atendido os critérios de qualidade e a validade discriminante do modelo, avaliaremos, na próxima subseção, o modelo estrutural.

Tabela 2. **Critérios de qualidade**

Variáveis	VME	Confiabilidade composta	Alfa de Cronbach	EC	ID	CGI	IA	EV
Expectativa de custo	0,778	0,875	0,715	**0,882**				
Intenção de desenvolvimento	0,698	0,874	0,784	-0,304	**0,836**			
CGI	0,548	0,923	0,907	-0,407	0,258	**0,740**		
Intenção de aquisição	0,657	0,851	0,747	-0,405	0,735	0,300	**0,811**	
Expectativa de valor	0,819	0,901	0,780	-0,392	0,318	0,647	0,360	**0,905**
Média				4,75	3,26	4,18	3,40	5,16
DP				1,64	1,87	1,64	1,92	1,67

Nota: VME = Variância Média Extraída; EC = Expectativa de Custo; ID = Intenção de Desenvolvimento; CGI = Capacidade de Gestão da Informação; IA = Intenção de Aquisição; EV = Expectativa de Valor.

Tabela 3. **Cargas cruzadas**

Itens x Variáveis	CGI	EC	ID	IA	EV
CGI1	**0,585**	-0,178	0,022	0,004	0,363
CGI2	**0,757**	-0,255	0,236	0,263	0,459
CGI3	**0,784**	-0,273	0,177	0,165	0,543
CGI4	**0,823**	-0,347	0,319	0,351	0,656
CGI5	**0,817**	-0,289	0,190	0,203	0,600
CGI6	**0,697**	-0,182	0,033	-0,048	0,349
CGI7	**0,735**	-0,265	0,308	0,480	0,486
CGI8	**0,686**	-0,293	0,107	0,286	0,425
CGI9	**0,711**	-0,417	0,125	0,191	0,337
CGI10	**0,773**	-0,455	0,259	0,186	0,452
EC1	-0,387	**0,885**	-0,299	-0,316	-0,390
EC2	-0,331	**0,879**	-0,237	-0,399	-0,301
ID1	0,253	-0,285	**0,826**	0,819	0,305
ID2	0,239	-0,204	**0,892**	0,588	0,253
ID3	0,145	-0,261	**0,786**	0,385	0,226
IA1	0,253	-0,285	0,826	**0,819**	0,305
IA2	0,249	-0,404	0,481	**0,858**	0,362
IA3	0,229	-0,269	0,517	**0,751**	0,166
EV1	0,557	-0,361	0,325	0,362	**0,907**
EV2	0,615	-0,349	0,250	0,289	**0,903**

Avaliação do modelo estrutural

Para testar as hipóteses e o poder preditivo do modelo, foram calculados os coeficientes de determinação de Pearson (R^2), o tamanho do efeito (f^2), a validade preditiva (Q^2) e o coeficiente de caminho (r). De acordo com os critérios de Cohen (1988), pode-se verificar um efeito médio do modelo sobre as variáveis expectativa de custo (EC) (0,166) e intenção de desenvolvimento (ID) (0,139), um grande efeito sobre a variável expectativa de valor (EV) (0,419) e um efeito relativamente grande sobre a variável intenção de aquisição (0,212).

A análise de *bootstrapping* com mil amostras demonstra que todas as relações das variáveis observáveis com as variáveis latentes, e aquelas entre as variáveis latentes, apresentam correlações significativas e coeficientes de regressão considerando p<0,001, o que faz com que Ho seja rejeitada. Em seguida, foram realizadas duas outras avaliações de qualidade do ajuste do modelo, a validade preditiva (Q^2) e o tamanho do efeito ($f2^2$), por meio do procedimento de olhos vendados. A Tabela 4 mostra que todos os Q^2s estão acima de zero, demonstrando a precisão do modelo. A análise do tamanho do efeito considera uma utilidade média de EC, ID e intenção de aquisição (IA) para ajustar o modelo. Os resultados aproximam-se de uma utilidade relativamente grande do EV de acordo com os critérios de Hair et al. (2016). Finalmente, os coeficientes de caminho, ilustrados na Figura 2, mostram que os resultados confirmam todas as hipóteses.

Tabela 4. **Resultados de R^2, Q^2 e f^2**

Relações	R^2	Q^2	f^2
EC	0,166	0,112	0,189
ID	0,139	0,085	0,143
IA	0,212	0,111	0,119
EV	0,419	0,333	0,339

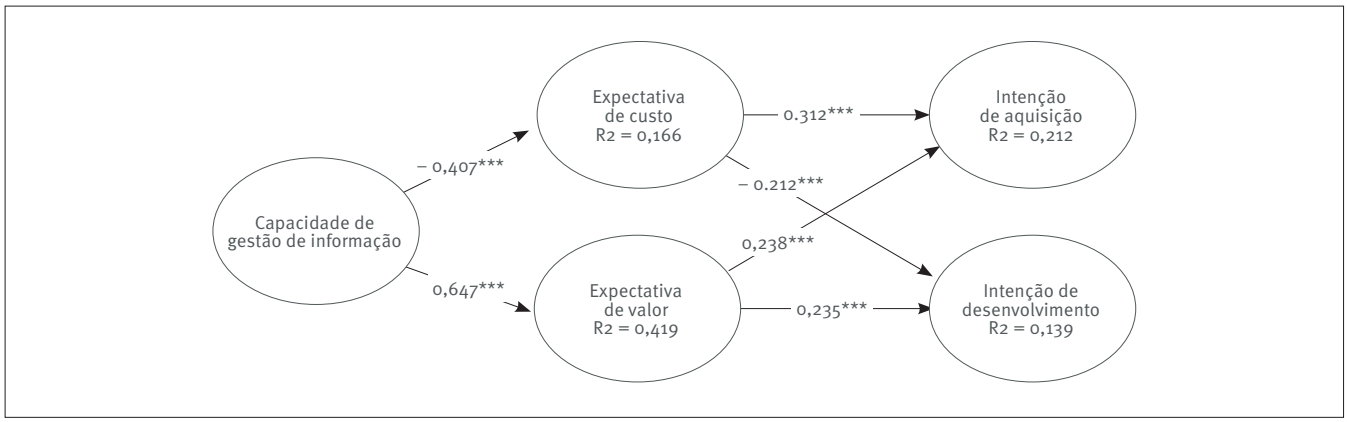
Considerando os pressupostos teóricos do MFE, H1 foi confirmado, uma vez que a CGI teve um impacto negativo sobre o CE de estratégias de *Big Data*; ou seja, quanto mais desenvolvido for a CGI da empresa, menor será a expectativa da despesa para implementar uma estratégia de *Big Data*. A análise do coeficiente de caminho destaca que o efeito da CGI é ainda mais evidente na expectativa do valor (EV) dessas estratégias. A Hipótese 2 foi confirmada, indicando que essa habilidade pode ser uma fonte potencial de imperfeições no MFE para *Big Data* em ambos os casos.

A outra metade do modelo (H3 e H4) descreve o impacto da expectativa de implementar estratégias de *Big Data* em termos de custo e valor na intenção de adquirir (H3 e H4) e desenvolver (H3b e H4b) essas capacidades. Ambas as hipóteses foram confirmadas. Esse impacto foi negativo para as Hipóteses 3a (compra) e 3b (desenvolvimento), demonstrando que uma alta expectativa de custo tem um impacto negativo na intenção de adquirir ou

 CC BY © RAE | São Paulo | **V. 59** | **n. 6** | **nov-dez 2019** | 379-388 ISSN 0034-7590; eISSN 2178-938X

desenvolver estratégias de *Big Data*. Os resultados também confirmaram as Hipóteses 4a e 4b. Em outras palavras, a intenção de adquirir ou desenvolver estratégias de *Big Data* foi positiva quando a expectativa de valor (EV) de uma estratégia de *Big Data* era maior.

Figura 2. **Resultados do modelo empírico: coeficientes de caminho e R²**



CONSIDERAÇÕES FINAIS

Finalizamos esta seção com uma discussão sobre o assunto e um esboço de futuras direções de pesquisa.

Contribuições para a literatura

Este artigo contribui para a literatura de sistemas de informação de gestão explorando um tema relativamente recente (*Big Data*) e sua relação com uma capacidade existente de uma empresa (CGI). Especificamente, analisamos esse fenômeno focando o seu impacto nas organizações. "Esse foco cria uma ligação mais estreita entre dados e modelos de negócios: nos preocupamos profundamente com a transformação de negócios e a criação de valor através de dados, e menos com algoritmos ou estruturas sem uma ligação com o valor empresarial" (Agarwal & Dhar, 2014, p. 445).

Em primeiro lugar, a pesquisa empregou uma teoria raramente utilizada em SI – MFE. Essa teoria, juntamente com a teoria do custo de transação (amplamente utilizada em SI), serviu como base para o desenvolvimento das hipóteses e confirmou a análise estatística. Considerando esse fundamento teórico e indicações encontradas na literatura, foi possível estabelecer a Hipótese 1. Nossos resultados atestam que a CGI pode ter um impacto negativo no custo esperado dos recursos necessários para implementar uma estratégia de *Big Data*. Esses resultados confirmam que as organizações têm expectativas de custo diferentes na busca de recursos estratégicos (Barney, 1986). A CGI desempenha um papel relevante nessa variedade de percepções, seja por meio de maior acurácia (Mithas et al., 2011) no acesso e distribuição da informação, seja pelo efeito perceptivo de polarização (Vasconcelos et al., 2006). As empresas que não conseguiram desenvolver com êxito a CGI podem ter uma expectativa mais elevada do custo para implementar uma nova estratégia relacionada com GI. Entretanto, esse efeito parece ser mais evidente nas relações da Hipótese 2. Demonstramos que a CGI impacta positivamente a extração de valor esperada de uma estratégia de *Big Data*. Esse foi o efeito mais significativo que encontramos, podendo indicar um produto das habilidades desenvolvidas ou um reflexo de experiências bem-sucedidas em GI.

Por outro lado, nas Hipóteses 3a (compra) e 3b (desenvolver), explicamos o impacto do custo esperado na intenção de aquisição ou desenvolvimento dos recursos e capacidades para implementar uma estratégia para lidar com dados volumosos e heterogêneos. O impacto negativo foi confirmado por dados empíricos, demonstrando que uma expectativa de custo elevada tem um impacto ainda mais negativo na intenção de aquisição do que na intenção de desenvolver os recursos e capacidades necessários para estabelecer a estratégia internamente. Por outro lado, os resultados confirmaram a Hipótese 4 (H4a e H4b), mostrando que uma maior expectativa de extração futura de valor impacta positivamente a intenção de adquirir ou desenvolver estratégias de *Big Data*. Nesse caso, os efeitos de tamanho encontrados para a intenção de adquirir ou desenvolver os recursos necessários para essas estratégias foram muito semelhantes. No entanto, este estudo não teve como objetivo avaliar se essas expectativas correspondem ou não à realidade do mercado. É importante notar que, no geral, investimentos em estratégias de SI só reduzem os custos de transação se a empresa consome menos recursos do que a quantidade gerada pela economia (Ciborra, 1996).

Através de duas perspectivas teóricas, nossa pesquisa contribui para a compreensão do impacto que a CGI existente pode ter na adoção ou não de novas estratégias em resposta a mudanças na gestão da informação. Ainda mais importante, este estudo revelou o papel dessa capacidade como fonte potencial de imperfeições no MFE, podendo, assim, ser um primeiro passo para investigar o papel da CGI no desempenho competitivo das empresas.

Além disso, juntamente com a adoção da perspectiva da literatura sobre CGI, propomos uma nova definição mais alinhada com o contexto atual e as necessidades de GI das organizações. Também propusemos e validamos uma nova escala para medir esse construto.

Implicações práticas

Podemos categorizar as implicações práticas deste estudo para dois tipos de organizações: aquelas que buscam soluções para responder às mudanças ambientais causadas pelo *Big Data* e aquelas que oferecem essas soluções. Para as empresas que planejam implementar estratégias de *Big Data*, os resultados revelam uma grande variação nas expectativas tanto do valor quanto do custo dos recursos necessários. Essa variação pode traduzir-se em oportunidades de buscar no mercado por recursos subestimados ou de incorrer no risco de adquirir recursos sobrevalorizados. Para reduzir esses riscos e melhorar o desempenho na exploração dessas oportunidades, nossos resultados mostram que investir em GI não só melhora o desempenho organizacional (Carmichael et al., 2011; Mithas et al., 2011), como também pode ajudar as empresas a avaliar estratégias futuras.

Já para o outro lado do mercado, este estudo pode prover, para empresas que oferecem os recursos e capacidades para implementar estratégias de *Big Data*, algum *insight* sobre as expectativas de seus consumidores atuais ou potenciais. Compreender as diferenças nas percepções de organizações com níveis diversos de CGI pode ajudar as empresas a criar uma solução adequada e contribuir para o sucesso dessa solução no desenvolvimento da CGI em níveis mais elevados para seus clientes.

Limitações e estudos futuros

Nossa amostra de estudo foi muito heterogênea, como mostra a Tabela 1, pois coletamos dados de maneira não sistemática, o que pode, portanto, não refletir inteiramente a população de empresas. Também não é possível identificar se os resultados se aplicam a um grupo específico de organizações. Mensuramos os construtos de intenção de aquisição e expectativa de custo utilizando apenas dois indicadores, e, apesar de ambos apresentarem bom desempenho em termos de validade e confiabilidade, ainda assim, utilizamos um indicador a menos do que o recomendado.

Esta pesquisa abre caminho para novas investigações em SI, em especial as relacionadas à CGI, o contexto do *Big Data*, e até mesmo novos estudos utilizando a teoria do MFE. Em relação à CGI, acreditamos que pesquisas futuras podem confirmar o papel estratégico dessas capacidades, especialmente nesse contexto de *Big Data*. Os pesquisadores podem usar o MFE para analisar outros fenômenos na área e conectá-lo a outras teorias na literatura de SI. O modelo pode mostrar-se válido para as estratégias de SI em geral e pode ser investigado no contexto de outras tecnologias (como análise de negócios ou inteligência de negócios).

CONCLUSÃO

Este estudo, apesar de trazer resultados quantitativos, é exploratório, dada a natureza do conteúdo analisado. Buscamos investigar como a CGI preexistente dentro das organizações afeta as expectativas e intenções dessas empresas de adotar uma nova estratégia de GI.

Nossos resultados oferecem *insights* sobre o efeito exercido nas relações entre CGI e custo e expectativa de valor futuro, além do impacto dessas expectativas na intenção de adquirir ou desenvolver os recursos necessários para implementar uma estratégia de *Big Data*. De modo geral, os resultados revelaram que a CGI influencia positivamente a expectativa de valor e influencia negativamente a expectativa de custo. A expectativa de valor impacta de maneira homogênea e positiva a intenção de adquirir ou desenvolver esses recursos. Finalmente, a expectativa de custo influencia negativamente a intenção de desenvolvimento e, de modo ainda mais contundente, a intenção de aquisição dos recursos e capacidades de *Big Data*.

Se um recurso fundamental para a sobrevivência nesse novo ambiente é a capacidade de obter acesso a mais informações sendo capaz de gerenciar esse fluxo de informação (Cordella, 2006), esta pesquisa contribui para a literatura de SI explorando o potencial da CGI nesse contexto de *Big Data*. Do ponto de vista acadêmico, este estudo testou uma teoria pouco utilizada na literatura, que os pesquisadores podem explorar ainda mais para analisar temas de SI. Por fim, esta pesquisa pode ajudar empresas que fornecem soluções de *Big Data*, bem como empresas que pretendem investir em estratégias para lidar com essa mudança no ambiente da gestão da informação.

AGRADECIMENTOS

Os autores agradecem o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

REFERÊNCIAS

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

FÓRUM

Submetido 26.09.2018. Aprovado 19.07.2019
Avaliado pelo sistema *double blind review*. Editores Científicos Convidados: Eduardo de Rezende Francisco, José Luiz Kugler, Soong Moon Kang, Ricardo Silva e Peter Alexander Whigham
Versão traduzida
DOI: http://dx.doi.org/10.1590/S0034-759020190605

INTENÇÃO DE ADOÇÃO DE *BIG DATA* NA CADEIA DE SUPRIMENTOS: UMA PERSPECTIVA BRASILEIRA

Intention to adopt big data in supply chain management: A Brazilian perspective

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

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, PLS-SEM.

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, PLS-SEM.*

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, PLS-SEM.*

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, Brasil

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

INTRODUÇÃO

O rápido avanço de tecnologias da informação e comunicação (TICs) motivou profissionais e estudiosos de logística e cadeia de suprimentos (Zinn & Goldsby, 2017a, 2017b) a entender o que as organizações podem agregar com essas tecnologias. Uma tecnologia inovadora e impactante que surgiu recentemente é o big data (grande volume de dados) (Davenport, 2006; Manyika et al., 2011; Rotella, 2012). A quantidade de dados produzidos diariamente tem aumentado drasticamente nos últimos anos (Domo, 2017). Esse crescimento trouxe diversos desafios ao gerenciamento de dados. Nesse contexto, o big data é uma abordagem robusta para ajudar organizações a analisarem (Croll, 2015) grandes quantidades de dados e aprimorarem o processo de tomada de decisão (Abawajy, 2015).

A literatura considera o big data a "próxima grande inovação" (Gobble, 2013, p. 64) e "o quarto paradigma da ciência" (Strawn, 2012, p. 34), e que tem impacto em praticamente todos os modelos de negócios. Por exemplo, 35% da receita da Amazon.com é gerada utilizando big data (Wills, 2014), juntamente com a reformulação de atividades de *marketing* que coletam dados importantes sobre o comportamento dos consumidores em tempo real (Erevelles, Fukawa, & Swayne, 2016). Um campo que vem realizando esforços significativos para utilizar o *big data* é a gestão da cadeia de suprimentos (GCS) (Gunasekaran et al., 2017; Kache & Seuring, 2017; Richey, Morgan, Lindsey-Hall, & Adams, 2016; Wu et al., 2017; Zhao, Liu, Zhang, & Huang, 2017).

Apesar dos potenciais benefícios do emprego do *big data* no GCS (Hazen, Boone, Ezell, & Jones-Farmer, 2014; Kache & Seuring, 2017; Schoenherr & Speier-Pero, 2015), ainda há pouco conhecimento sobre *big data* e poucas iniciativas para adotá-lo no mercado brasileiro de GCS, e a literatura carece de fortes resultados empíricos (Queiroz & Telles, 2018). O estágio atual de utilização de *big data* constitui uma oportunidade para acadêmicos e profissionais preencherem essa lacuna. Por exemplo, até onde sabemos, nenhum estudo anterior analisou a intenção comportamental (IC) dos profissionais brasileiros de GCS de adotar *big data*, e o presente estudo visa preencher essa lacuna. Adaptamos um modelo previamente desenvolvido e validado de aceitação e uso de tecnologia (*Unified theory of acceptance and use of technology* - [UTAUT]) (Venkatesh, Morris, Davis & Davis, 2003; Queiroz & Wamba, 2019), incluindo um construto de confiança. Mais especificamente, este estudo responde à seguinte pergunta: Como as variáveis do modelo UTAUT explicam a IC dos profissionais brasileiros de GCS de adotar big data?

Para responder a essa pergunta, este trabalho baseia-se na literatura relacionada ao *big data* (Davenport, 2006; Manyika et al., 2011; Queiroz & Telles, 2018), GCS (Carter, Rogers, & Choi, 2015; Mentzer et al., 2001) e UTAUT (Queiroz & Wamba, 2019; Venkatesh et al., 2003; Venkatesh, Thong, & Xu, 2012) para desenvolver as hipóteses e o modelo. O modelo conceitual foi adaptado e validado utilizando modelos de equações estruturais de mínimos quadrados parciais (*partial least squares structural equation modeling* – PLS-SEM). As principais conclusões têm fortes implicações teóricas e gerenciais. Da perspectiva gerencial, verificamos que condições facilitadoras (exemplo: infraestrutura) exercem grande influência sobre a IC de adotar *big data*. Sob a ótica teórica, nossos resultados revelaram que a influência social e expectativa de desempenho não são bons preditores da IC de profissionais brasileiros de GCS em adotar *big data*.

Este artigo está organizado da seguinte forma. Inicialmente, apresentamos os principais fundamentos teóricos para estudos de *big data*, GCS e UTAUT. Em seguida, descrevemos as hipóteses, o modelo de pesquisa e metodologia, bem como a análise baseada em PLS-SEM. Discutimos as implicações gerenciais e teóricas, assim como as limitações do presente estudo e orientações para pesquisas futuras. Finalmente, as principais conclusões são destacadas.

BASES TEÓRICAS

Big data: Fundamentos, conceitos e desafios

O *big data* surgiu como uma TIC altamente inovadora. Uma definição abrangente e adequada de *big data* é "[...] conjuntos de dados cujo tamanho está além da capacidade das ferramentas de software de banco de dados de coletar, armazenar, gerenciar, e analisar" (Manyika et al., 2011, p. 1). Assim, o *big data* pode ser considerado uma abordagem robusta para analisar dados no contexto de decisões descritivas, prescritivas e preditivas (Phillips-Wren & Hoskisson, 2015). Essa abordagem é comumente chamada de *big data analytics* (BDA) e envolve uma abordagem de cinco Vs (volume, velocidade, variedade, veracidade e valor) (Queiroz & Telles, 2018; Wamba et al., 2017). Em outras palavras, o BDA usa estatísticas sofisticadas e modelos matemáticos e computacionais para analisar grandes volumes de dados e fornecer informações aos tomadores de decisão. Neste estudo, usamos a definição de *big data* proposta por Phillips-Wren e Hoskisson (2015) como a quantidade de dados que supera a capacidade das organizações de armazená-los e analisá-los para apoiar e melhorar o processo de tomada de decisão.

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

O volume de dados aumentou drasticamente nos últimos anos devido à variedade de dados produzidos (Bibri & Krogstie, 2017) (por exemplo, ERP, Twitter, Facebook, Google, LinkedIn, GPS, entre outros) e à velocidade da transmissão de dados (Munshi & Mohamed, 2017; Srinivasan & Swink, 2018). Esse cenário complexo impulsiona organizações a desenvolverem recursos específicos para armazenar, processar e analisar dados para auxiliar o processo de tomada de decisão. No entanto, criar valor não é uma tarefa fácil, principalmente porque a capacidade das organizações de processar e analisar dados é limitada. Além disso, confirmar a veracidade dos dados, que indica a sua qualidade e confiabilidade (Munshi & Mohamed, 2017; Nobre & Tavares, 2017), é um grande desafio para as empresas.

Big data está sendo utilizado no aprimoramento do GCS utilizando sistemas baseados em múltiplos agentes (Giannakis & Louis, 2016), otimização do GCS verde considerando materiais perigosos e emissão de carbono (Zhao et al., 2017), setor manufatureiro (Zhong, Newman, Huang, & Lan, 2016) e exploração de informações de GCS (Kache & Seuring, 2017). Foi demonstrado que o *big data* pode melhorar significativamente o desempenho das empresas (Akter, Wamba, Gunasekaran, Dubey, & Childe, 2016; Gunasekaran et al., 2017; Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015; Wang, Gunasekaran, Ngai, & Papadopoulos, 2016).

Gestão da cadeia de suprimentos e os impactos de tecnologias de ponta

Recentemente, os campos de logística e GCS foram impactados significativamente pelo crescimento exponencial do uso de TICs. Dessa forma, acadêmicos e empresários procuraram entender os possíveis efeitos e oportunidades de aplicar o GCS em seus modelos de negócios (Zinn & Goldsby, 2017a, 2017b). Nesse contexto, o GCS é definido como:

> O gerenciamento de uma rede de relacionamentos dentro de uma empresa e entre organizações e unidades de negócios interdependentes, incluindo fornecimento de materiais, compras, produção, logística, marketing, e sistemas relacionados que facilitam o fluxo de materiais, serviços, finanças, e informações do produtor original ao cliente final com o objetivo de agregar valor, maximizar a lucratividade por meio de eficiências, e satisfazer o cliente. (Stock & Boyer, 2009, p. 706)

Além disso, o GCS pode ser considerado uma rede (Carter et al., 2015) e um sistema adaptativo complexo (Choi, Dooley e Rungtusanatham, 2001), e a crescente quantidade de dados tem aumentado essa complexidade. O uso de big data no GCS auxilia o processo de tomada de decisão, fornecendo informações relevantes sobre a dinâmica do sistema (por exemplo, padrões de compra de clientes, análise de custos e tendências de mercado). Análises descritivas e prescritivas robustas (Wang et al., 2016) têm melhorado significativamente o desempenho de empresas (Akter et al., 2016; Gunasekaran et al., 2017).

Modelos de aceitação de tecnologia (MATs) e Teoria unificada de aceitação e uso de tecnologia (UTAUT)

Vários estudos avaliaram o desenvolvimento e disseminação da tecnologia da informação (TI) (Davis, 1989; Morris & Venkatesh, 2000; Venkatesh & Brown, 2001; Wamba, 2018) e valores e comportamentos individuais em relação à adoção e uso de TI (Mamonov & Benbunan-Fich, 2017; Youngberg, Olsen, & Hauser, 2009). O MAT é uma contribuição importante na adoção da tecnologia (Davis, 1989) e tem suas raízes na teoria da ação racional (Azjen & Fishbein, 1980). Os fundamentos do MAT baseiam-se em duas variáveis: utilidade percebida e facilidade de uso percebida. Mais recentemente, Venkatesh et al. (2003) propuseram a consolidação das teorias do MAT que antecederam a UTAUT.

UTAUT

O modelo UTAUT (Venkatesh et al., 2003) é uma abordagem robusta e relevante para entender comportamentos individuais de adoção e uso da tecnologia. O modelo possui quatro constructos voltados a intenção de uso da tecnologia: expectativa de desempenho (ED), expectativa de esforço (EE), influência social (IS) e condições facilitadoras (CFs).

A ED refere-se ao "grau em que um indivíduo acredita que o uso do sistema o ajudará a melhorar seu desempenho profissional" (Venkatesh et al., 2003, p. 447). A EE é "o grau de facilidade associado ao uso do sistema" (Venkatesh et al., 2003, p. 450). A IS denota "o grau em que um indivíduo percebe que outras pessoas importantes acreditam que ele ou ela deve usar o novo sistema" (Venkatesh et al., 2003, p. 451). As CFs indicam "o grau em que um indivíduo acredita que existe uma infraestrutura organizacional e técnica de suporte ao uso do sistema" (Venkatesh et al., 2003, p. 453). O modelo UTAUT possui quatro moderadores: gênero, idade, experiência e voluntariedade de uso. No entanto, seguindo as

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

premissas de um estudo anterior (Weerakkody, El-Haddadeh, Al-Sobhi, Shareef, & Dwivedi, 2013), esses moderadores não foram empregados no presente modelo adaptado (explicado na próxima seção) porque se trata de um estudo preliminar de adoção de BDA no contexto brasileiro de GCS.

Hipóteses e modelo de pesquisa

Com o apoio da literatura relevante sobre *big data*, GCS e UTAUT, adaptamos um modelo descrito por Queiroz e Wamba (2019) para compreender o comportamento de adoção de *big data* por profissionais brasileiros de GCS. Adotamos alguns conceitos e hipóteses propostos no modelo de Queiroz e Wamba (2019) (Figura 1), pois estes foram previamente validados (Quadro 1), e um constructo de confiança previamente validado que foi adicionado a esses conceitos (Alalwan, Dwivedi, & Rana, 2017; Gefen, Karahanna, & Straub, 2003). Além disso, os constructos utilizados em nosso modelo têm relacionamentos diferentes daqueles descritos na literatura (Queiroz & Wamba, 2019).

Condições facilitadoras

As CFs desempenham um papel fundamental na previsão da aceitação e intenção de uso de tecnologia (Venkatesh et al., 2003, 2012). Neste estudo, as CFs denotam o conhecimento dos profissionais de GCS sobre os recursos e infraestrutura organizacional disponíveis para dar suporte ao uso de *big data*. Estudos anteriores relataram que CFs são um bom preditor da IC de adotar *big data* (Huang, Liu, & Chang, 2012; Sabi, Uzoka, Langmia, & Njeh, 2016). O presente estudo postula que CFs influenciam diretamente a IC e são fundamentais na EE (Dwivedi et al., 2017) e também influenciam a ED (Wang, Jeng & Huang, 2017). Portanto, propomos as seguintes hipóteses:

H1a: CFs afetam positivamente a EE.

H1b: CFs afetam positivamente a ED.

H1c: CFs afetam positivamente a IC de adotar *big data*.

Confiança

O conceito de confiança foi estudado exaustivamente nos campos de administração e sistemas de informação (Colquitt & Rodell, 2011; Wu, Zhao, Zhu, Tan, & Zheng, 2011). Confiança é definida como "a predisposição de uma pessoa de se tornar vulnerável às ações de outra pessoa (contraparte) com base na expectativa de que a contraparte realize uma ação específica importante para o credor, independentemente da capacidade de monitorar ou controlar a contraparte" (Mayer, Davis, & Schoorman, 1995, p. 712). Essa definição indica que confiança é a disposição de depender de uma contraparte com base na integridade, benevolência e credibilidade. Nesse contexto, o *big data* é confiável para os usuários. Em consonância com trabalhos anteriores (Wu et al., 2011), propomos as seguintes hipóteses:

H2a: Confiança afeta positivamente a ED.

H2b: Confiança afeta positivamente a IC de adotar *big data*.

Influência social

A IS é um bom preditor da IC de usar a tecnologia (Venkatesh et al., 2003). Neste trabalho, a IS denota até que ponto os profissionais de GCS acreditam que seus colegas devem usar *big data*. Estudos anteriores destacam que a IS afeta a IC (Batara, Nurmandi, Warsito, & Pribadi, 2017; Oliveira, Faria, Thomas, & Popovič, 2014; Venkatesh et al., 2012). A nossa hipótese é que, no contexto do GCS, a IS afeta significativamente a confiança (Chin, Wafa, & Ooi, 2009) e, por sua vez, a IC (Alalwan et al., 2017). Portanto, propomos as seguintes hipóteses:

H3a: IS afeta positivamente a confiança.

H3b: IS afeta positivamente a IC de adotar *big data*.

Expectativa de esforço

A EE está relacionada à complexidade de operação do sistema (Venkatesh et al., 2003). Neste estudo, a EE refere-se à facilidade de uso de sistemas de *big data* por profissionais de GCS. Estudos anteriores discutiram o efeito da EE na IC e uso de uma nova tecnologia (Batara et al., 2017; Venkatesh et al., 2012; Zhao, Ni, & Zhou, 2018) e na adoção de *blockchain* em GCS (Francisco & Swanson, 2018). Nesse sentido, a seguinte hipótese foi postulada:

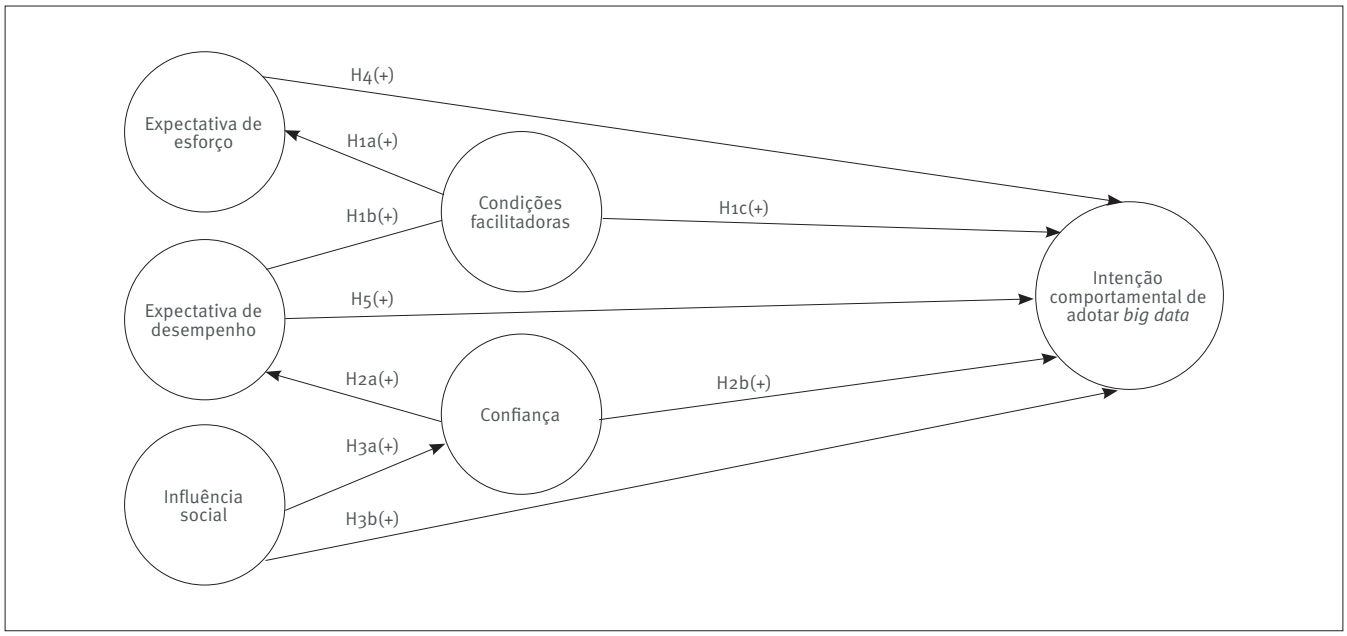
H4: A EE afeta positivamente a IC de adotar *big data*.

Expectativa de desempenho

Neste trabalho, a ED denota o nível em que um profissional de GCS percebe que o *big data* melhorará sua produtividade e desempenho. As empresas podem analisar diferentes tipos de dados utilizando estatísticas robustas e técnicas de aprendizado de máquina (Kune, Konugurthi, Agarwal, Chillarige, & Buyya, 2016), resultando em considerável economia de tempo e melhora da produtividade nas organizações, aumentando seu desempenho (Gunasekaran et al., 2017; Wamba et al., 2017). Portanto, apresentamos a seguinte hipótese:

H5: ED afeta positivamente a IC de adotar *big data*.

Figura 1. **Modelo conceitual**



METODOLOGIA

Amostra e coleta de dados

Um questionário baseado em Queiroz e Wamba (2019) foi utilizado para testar as hipóteses propostas. O questionário *on-line* foi baseado em conceitos e escalas validados anteriormente (Venkatesh et al., 2003, 2012; Gefen et al., 2003). O modelo de Queiroz e Wamba (2019) foi desenvolvido com base em estudos anteriores, e seus constructos foram adaptados de estudos recentes envolvendo MATs (Alalwan et al., 2017; Venkatesh et al., 2003, 2012). Utilizamos o questionário desenvolvido por Queiroz e Wamba (2019) porque nosso principal objetivo era identificar a intenção de adotar *big data*. Todas os constructos foram mensurados usando uma escala *Likert* de sete pontos, variando de 1 [discordo totalmente] a 7 [concordo totalmente] (Wamba et al., 2017). Antes da coleta de dados, foi realizado um teste-piloto com cinco acadêmicos seniores e cinco profissionais seniores de GCS. Os dados foram obtidos da rede social *LinkedIn* (Gupta & George, 2016; Queiroz & Telles, 2018). Após o teste-piloto, enviamos um questionário a 600 profissionais brasileiros de GCS com experiência em *big data*. A pesquisa foi realizada em agosto de 2018, em que recebemos 152 questionários válidos, representando uma taxa de resposta de 25,33%. Os constructos e seus respectivos itens estão mostrados no Quadro 1. O questionário foi validado empregando cargas fatoriais (Hair, Hult, Ringle, & Sarstedt, 2017), alfa de Cronbach, confiabilidade composta, variância média extraída (Hair et al., 2017; Nunnally, 1978; Riffai, Grant, & Edgar, 2012) e validade discriminante.

RESULTADOS E ANÁLISES

O PLS-SEM (Ringle, Wende, & Becker, 2015; Shim, Lee, & Kim, 2018; Sun & Teng, 2017) foi utilizado para analisar o modelo de estudo. O PLS-SEM é uma abordagem eficaz para analisar modelos simples e robustos da área de gestão e negócios (Hair, Sarstedt, Hopkins, & Kuppelwieser, 2014; Hair et al., 2017) e ganhou a atenção de estudiosos de GCS (Autry, Williams, & Golicic, 2014; Grawe, Daugherty, & Ralston, 2015; Han, Wang, & Naim, 2017; Yadlapalli, Rahman, & Gunasekaran, 2018). Suas principais vantagens são a flexibilidade no trabalho com amostras pequenas (por exemplo, 100 respondentes) e a presença de constructos formativos e reflexivos (Hair et al., 2017).

As características dos entrevistados estão apresentadas na Tabela 1. Os entrevistados do sexo masculino representaram cerca de 90% da amostra. A maioria (52,63%) dos entrevistados tinha entre 34 e 41 anos. Os entrevistados possuíam pós-graduação/MBA (nível acadêmico mais alto da amostra, 55,26%), diploma de bacharel (39,47%), ou mestrado acadêmico (5,26%). O tempo de experiência profissional dos entrevistados em suas respectivas organizações foi de 2 a 5 anos (50%), de 6 a 10 anos (21,05%), ou menos de 1 ano (18,42%). A amostra foi composta por analistas de logística (46,05%), gerentes de transporte (26,32%), gerentes de operações (18,42%) e gerentes de GCS (9,21%).

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

Quadro 1. **Itens de mensuração**

Constructo	Descrição	Itens	Fontes
Expectativa de desempenho (ED)	ED1	*Big data* é útil na minha vida diária.	(Alalwan et al., 2017; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019)
	ED2	O uso de *big data* aumenta minhas chances de concluir tarefas importantes para mim.	
	ED3	O emprego de *big data* me ajuda a realizar tarefas mais rapidamente.	
	ED4	A aplicação de *big data* aumenta minha produtividade.	
Expectativa de esforço (EE)	EE1	Aprender a utilizar *big data* é fácil para mim.	(Alalwan et al., 2017; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019)
	EE2	Minha interação com o *big data* é clara é adequada.	
	EE3	*Big data* é fácil de usar.	
	EE4	Consigo me tornar proficiente no uso de *big data* rapidamente.	
Influência social (IS)	IS1	As pessoas que são importantes para mim pensam que eu deveria empregar *big data*.	(Alalwan et al., 2017; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019)
	IS2	As pessoas que são importantes para mim acham que eu deveria utilizar *big data*.	
	IS3	Pessoas cujas opiniões eu valorizo preferem que eu use *big data*.	
Condições facilitadoras (CFs)	CF1	Eu tenho os recursos necessários para adotar o *big data*.	(Alalwan et al., 2017; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019)
	CF2	Eu tenho o conhecimento necessário para adotar *big data*.	
	CF3	*Big data* é compatível com outras tecnologias que eu utilizo.	
	CF4	Posso obter ajuda de outras pessoas quando tenho dificuldades em usar *big data*.	
Intenção comportamental (IC) de utilizar *big data*	IC1	Pretendo usar *big data* no futuro.	(Alalwan et al., 2017; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019)
	IC2	Espero utilizar *big data* no futuro.	
	IC3	Tenho planos para adotar *big data* no futuro.	
Confiança (CO)	CO1	Acredito que *big data* é confiável.	(Alalwan et al., 2017; Gefen et al., 2003)
	CO2	Confio em *big data*.	
	CO3	Não duvido da credibilidade do *big data*.	
	CO4	Tenho certeza de que estruturas legais e tecnológicas me protegem adequadamente de problemas em *big data*.	
	CO5	*Big data* cumpre as funções pretendidas.	

Tabela 1. **Perfil demográfico da amostra (n=152)**

Gênero	n	%
Masculino	136	89,5
Feminino	16	10,5
Idade (anos)		
26-33	40	26,32
34-41	80	52,63
42-49	12	7,89
≥50	20	13,16
Nível educacional		
Bacharelado	60	39,47
Pós-graduação/MBA	84	55,26
Mestrado acadêmico	8	5.26
Tempo de experiência profissional (anos)		
‹1	28	18,42
2-5	76	50,00
6-10	32	21,05
11-15	16	10,53
Profissão		
Analista de logística	70	46,05
Gerente de operações	28	18,42
Gerente de transporte	40	26,32
Gerente de cadeia de suprimentos	14	9,21

O modelo de estudo foi analisado utilizando o *software* SmartPLS versão 3.0 (Hair et al., 2017; Ringle et al., 2015). Primeiro, o modelo foi avaliado por suas cargas fatoriais, alfa de Cronbach, confiabilidade composta, variância média extraída e validade discriminante.

Modelo de mensuração

Todas as cargas fatoriais excederam o limiar de 0,70 recomendado na literatura (Hair et al., 2017) (Tabela 2). As principais medidas de confiabilidade do constructo e consistência interna dos itens são mostradas na Tabela 3. O valor alfa de Cronbach e a confiabilidade composta ficaram acima do limite de 0,70, e todos os valores médios de variância extraídos ficaram acima do limiar de 0,50 (Hair et al., 2017; Nunnally, 1978; Riffai et al., 2012). Portanto, todos os constructos adotados no modelo foram validados. Os resultados de validade discriminante estão descritos na Tabela 4. Nesse caso, a raiz quadrada da variância média extraída de cada constructo deve ser maior que as correlações entre os constructos (Fornell & Larcker, 1981; Henseler, Ringle, & Sinkovics, 2009). Os valores obtidos foram superiores ao limiar de 0,70 (Fornell & Larcker, 1981), demonstrando que todos os constructos apresentaram validade discriminante (Ahmad & Khalid, 2017; Martins, Oliveira, & Popovič, 2014).

Tabela 2. **Cargas fatoriais**

	IC	EE	CF	ED	IS	CO
IC1	0,887					
IC2	0,916					
IC3	0,900					
EE1		0,945				
EE2		0,906				
EE3		0,947				
EE4		0,900				
CF1			0,789			
CF2			0,884			
CF3			0,719			
CF4			0,800			
ED1				0,764		
ED2				0,803		
ED3				0,914		
ED4				0,899		
IS1					0,953	
IS2					0,983	
IS3					0,967	
CO1						0,964
CO2						0,944
CO3						0,918
CO4						0,937
CO5						0,913

IC, intenção comportamental;
EE, expectativa de esforço;
CF, condições facilitadoras;
ED, expectativa de desempenho;
IS, influência social; CO, confiança.

CC BY
© RAE | São Paulo | **V. 59** | **n. 6** | **nov-dez 2019** | 389-401 ISSN 0034-7590; eISSN 2178-938X

Tabela 3. **Medidas de confiabilidade**

Constructo	Alfa de Cronbach	Confiabilidade composta	Variância média extraída
IC	0,881	0,926	0,808
EE	0,942	0,959	0,853
CF	0,806	0,873	0,637
ED	0,864	0,908	0,715
IS	0,965	0,977	0,934
CO	0,963	0,971	0,871

Tabela 4. **Validade discriminante**

Constructo	IC	EE	CF	ED	IS	CO
IC	**0,901**					
EE	0,527	**0,925**				
CF	0,614	0,577	**0,800**			
ED	0,314	0,238	0,620	**0,847**		
IS	0,399	0,390	0,508	0,421	**0,968**	
CO	0,511	0,456	0,580	0,640	0,715	**0,935**

IC, intenção comportamental; EE, expectativa de esforço; CF, condições facilitadoras; ED, expectativa de desempenho; IS, influência social; CO, confiança.

Modelo estrutural

Os resultados do modelo estrutural estão nas Tabelas 5 e 6. A Tabela 5 destaca as estatísticas dos coeficientes de correlação. Os resultados corroboraram as seguintes hipóteses: H1a (CFs têm um efeito positivo significativo na EE) (β = 0,578; p ‹ 0,001), H1b (CFs têm um efeito positivo significativo na ED) (β = 0,380; p ‹ 0,001), H1c (CFs têm um efeito positivo significativo na IC) (β = 0,490; p ‹ 0,001), H2a (a confiança tem um efeito positivo significativo na ED) (β = 0,413; p ‹ 0,001), H2b (a confiança tem um efeito positivo significativo na IC) (β = 0,327; p ‹ 0,05), e H3a (a IS tem um efeito positivo significativo na confiança) (β = 0,710; p ‹ 0,001). Os resultados considerando as outras hipóteses foram inesperados: H3b (a IS tem um efeito positivo significativo na IC). O relacionamento foi negativo e não-significante (β = -0,073; p = 0,519). Assim, H3b não foi suportada. H4 (o EE tem um efeito positivo significativo na IC) não foi suportada também (β = 0,166; p ‹ 0,1), assim como H5 (a ED tem um efeito positivo significativo na IC) (β = -0,214; p ‹ 0,05). A Tabela 6 destaca as seguintes variâncias que foram obtidas: IC (46,0%), EE (33,3%), ED (49,8%) e confiança (50,3%) (Tabela 6). Todos os valores de R2 do modelo excederam o limiar de 0,20, o que corrobora estudos anteriores (Chin, 1998; Martins et al., 2014).

Tabela 5. **Coeficientes de caminhos**

Valor	Beta	Desvio-padrão	Estatísticas t	Valor de p	Resultados
CF -› EE	0,578	0,053	10,921*	0,000	Aceito
CF -› ED	0,380	0,064	5,875*	0,000	Aceito
CF -› IC	0,490	0,097	5,016*	0,000	Aceito
CO -› ED	0,413	0,080	5,301*	0,000	Aceito
CO -› IC	0,327	0,112	2,987**	0,003	Aceito
IS -› CO	0,710	0,047	15,17*	0,000	Aceito
IS -› IC	–0,073	0,097	0,646	0,519	Rejeitado
EE -› IC	0,166	0,090	1,86	0,063	Rejeitado
ED -› IC	–0,214	0,099	2,184**	0,029	Rejeitado

IC, intenção comportamental; EE, expectativa de esforço; CF, condições facilitadoras; ED, expectativa de desempenho; IS, influência social; CO, confiança.
*p ‹ 0,001; **p ‹ 0,05. .

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

Tabela 6. **Resultados de R^2 (variáveis dependentes)**

Constructo	R^2	R^2 ajustado
IC	0,477	0,460
EE	0,337	0,333
ED	0,504	0,498
CO	0,506	0,503

IC, intenção comportamental; EE, expectativa de esforço; ED, expectativa de desempenho; CO, confiança

DISCUSSÃO E IMPLICAÇÕES

O objetivo deste estudo foi compreender a intenção de adotar *big data* no contexto do GCS no Brasil. Em vista da falta de estudos sobre tecnologias de ponta no Brasil (Queiroz e Telles, 2018), este trabalho contribui para o conhecimento sobre BDA, GCS e MATs. Os resultados têm implicações gerenciais e teóricas significativas e oferecem orientações valiosas para adaptar e melhorar o modelo adotado.

Implicações gerenciais

As principais conclusões deste estudo têm implicações importantes para gerentes e profissionais de GCS envolvidos com BDA e seus facilitadores de adoção. De acordo com a literatura que considera o big data uma ferramenta essencial para melhorar o GCS (Gunasekaran et al., 2017; Hazen, Skipper, Ezell, & Boone, 2016; Wang et al., 2016), o presente estudo demonstrou que o big data pode ser uma ferramenta adequada para ajudar os profissionais que trabalham no GCS a obterem conhecimento e tomarem decisões. Segundo, as CFs exercem uma grande influência na adoção de *big data*. Isso significa que os gerentes devem considerar a infraestrutura de TI, velocidade da internet e integração com outros sistemas, além de outros fatores (Sabi et al., 2016; Venkatesh et al., 2003).

Surpreendentemente, apesar de a literatura indicar que a ED é um bom preditor da IC de adotar tecnologia (Dwivedi et al., 2017; Farooq et al., 2017; Venkatesh et al., 2003; Weerakkody et al., 2013), o presente estudo demonstrou que a ED não foi um bom preditor da IC de usar *big data* entre profissionais brasileiros de GCS. Esse resultado pode representar uma barreira à adoção desse tipo de tecnologia e abre novas oportunidades de pesquisa. Por outro lado, a IS é forte preditora de confiança (Chin et al., 2009). Entretanto, a IS não afetou a IC de adotar *big data*. Portanto, mais pesquisas que auxiliem tomadores de decisão são necessárias.

Implicações teóricas, limitações e perspectivas futuras

Do ponto de vista teórico, este estudo traz contribuições significativas para o campo da logística e da GCS. Primeiro, ao analisar a literatura sobre BDA, GCS e UTAUT, validamos um modelo teórico robusto. Adaptamos e aplicamos um modelo previamente desenvolvido, para utilização junto aos profissionais brasileiros de GCS, e nossos resultados validaram o modelo. O modelo teórico explicou 46% da IC, excedendo o limiar de 20% reportado na literatura (Chin, 1998; Martins et al., 2014). Além disso, os resultados de validade discriminante foram consistentes com a literatura e, portanto, apoiaram as hipóteses estruturais, e demonstraram que o modelo efetivamente mensura a IC de profissionais de GCS de adotar *big data*.

Nossos achados destacam que CFs são um bom preditor da IC de usar *big data*, e pesquisas futuras podem se concentrar em avaliar CFs e suas limitações. No modelo proposto, em consonância com um estudo anterior (Alalwan et al., 2017), a confiança foi um bom preditor de ED e IC. Por outro lado, a IS não foi uma boa preditora de IC, o que corrobora o resultado de um estudo anterior (Alalwan et al., 2017). Esse resultado sugere a necessidade de uma investigação mais aprofundada do papel da IS na implementação de *big data* e outras tecnologias emergentes na área de GCS.

Este estudo tem algumas limitações. Primeiro, uma variável moderadora poderia ser incorporada ao modelo (Venkatesh et al., 2003, 2012) para identificar nuances e diferenças na amostra, como tipo de indústria, gênero e experiência profissional. Segundo, devido a escassez de estudos sobre a adoção de *big data* no Brasil, nossos resultados possuem limitação relativo a comparação com outros estudos. No entanto, este estudo abre oportunidades para estudiosos e profissionais aplicarem o modelo validado e adaptarem-no a outros contextos. Terceiro, o modelo adotado foi testado em uma economia emergente; por esse motivo, os resultados não podem ser generalizados. Portanto, estudos adicionais são necessários para obter evidências empíricas em outros países utilizando esse modelo.

Este estudo é um dos primeiros a avaliar a IC de profissionais brasileiros de GCS de adotar *big data*. Há uma necessidade urgente e oportunidades de realizar investigações adicionais sobre essa e outras tecnologias de ponta (e.g., *blockchain*, internet das coisas e impressão 3D) e comparar as hipóteses desse modelo em outros contextos.

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

CONCLUSÃO

Os objetivos deste estudo foram: 1. elucidar a IC de profissionais brasileiros de GCS de empregar *big data*; e 2. aplicar um modelo contendo constructos relacionados à IC. Este estudo forneceu uma compreensão mais aprofundada da intenção de adotar o BDA no GCS no Brasil.

As três contribuições deste estudo foram: 1. com base em literatura teórica robusta (Akter et al., 2016; Alalwan et al., 2017; Davis, 1989; Venkatesh et al., 2003, 2012; Queiroz & Wamba, 2019), adaptamos e aplicamos um modelo para entender a IC de profissionais brasileiros de GCS; 2. os resultados têm fortes implicações teóricas e práticas; uma implicação teórica é que CFs e confiança são bons preditores de IC; em contrapartida, a IS não foi um preditor de IC, o que contraria os resultados de dois estudos (Venkatesh et al., 2003, 2012), mas corrobora um estudo (Alalwan et al., 2017); e 3. a ED e EE não são bons preditores de IC. Portanto, esses resultados abrem oportunidades para melhor entender esses dois constructos, ajuda a preencher uma lacuna na literatura empírica brasileira sobre o uso de *big data* no GCS e estimula profissionais de logística e GCS a melhor compreender essa tecnologia.

REFERÊNCIAS

Abawajy, J. (2015). Comprehensive analysis of big data variety landscape. International journal of parallel. *Emergent and Distributed 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

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 Busienss 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

FÓRUM

Submetido 19.10.2018. Aprovado 19.07.2019
Avaliado pelo sistema *double blind review*. Editores Científicos Convidados: Eduardo de Rezende Francisco, José Luiz Kugler, Soong Moon Kang, Ricardo Silva e Peter Alexander Whigham
Versão traduzida
DOI: http://dx.doi.org/10.1590/S0034-759020190606

MEDINDO A ACESSIBILIDADE: UMA PERSPECTIVA DE *BIG DATA* SOBRE OS TEMPOS DE ESPERA DO SERVIÇO DA UBER

Measuring accessibility: a Big Data perspective on uber service waiting times

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

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 de 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 múltipla (RLM), e, utilizando o método stepwise, foram selecionadas as variáveis mais significativas do modelo. Foi verificado o padrão espacial das variáveis por meio do teste I de Moran, que motivou a elaboração de um modelo espacial autorregressivo (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.

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.*

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, Brasil

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

INTRODUÇÃO

A introdução de empresas "ride-sourcing" (empresas de terceirização de viagens) no mercado privado de mobilidade urbana mudou consideravelmente os hábitos de muitos moradores das cidades, inclusive dos seus utilizadores tradicionais, os proprietários de automóveis particulares e passageiros.

O aumento da concorrência num mercado que tradicionalmente é altamente regulamentado levou a conflitos entre os novos serviços, os antigos fornecedores desse mercado, os táxis e as autoridades locais em muitas cidades. As dificuldades enfrentadas pelas autoridades para enquadrar esses novos serviços no quadro jurídico existente são claramente ilustradas no caso de São Francisco (Flores & Rayle, 2017).

Nesse contexto, empresas de terceirização de viagens tentaram construir uma imagem de serem amigas do ambiente, vendendo-se como serviços de compartilhamento de viagens (Flores & Rayle, 2017). Além disso, elas reivindicaram a capacidade de reduzir o número de veículos da rua (e suas emissões) e de oferecer serviços melhores e mais baratos a áreas anteriormente negligenciadas pelos táxis.

Nesse contexto, vários estudos procuraram esclarecer os efeitos das transformações resultantes. Jin, Kong, Wu e Sui (2018) revisaram estudos recentes sobre os efeitos dos serviços de terceirização de viagens. Entre os assuntos discutidos, estão a eficiência econômica e a equidade social desses serviços. Em termos gerais, parece haver evidência do aumento dos serviços de transporte privado em áreas periféricas e de baixa renda anteriormente negligenciadas pelos táxis. Apesar disso, o perfil médio do cliente do serviço é mais jovem, mais rico e mais educado do que o perfil médio da população; o artigo aponta para uma possível exclusão do serviço ao estilo de "divisão digital", tanto em termos de geração como de renda (Jin et al., 2018).

No entanto, outro aspecto essencial dos serviços de terceirização de viagens é a sua capacidade de gerar *Big Data* de alta qualidade que pode ser usado para avaliar várias questões relativas ao serviço. O acesso a esses dados permite uma análise rica e detalhada e pode revelar ainda mais sobre a interação entre outros modos de transporte e terceirização de viagens (Jin et al., 2018).

Em especial, a possibilidade de analisar a acessibilidade urbana a partir dessa fonte de dados é interessante. Como componente essencial da acessibilidade, a distribuição da rede de transportes (Páez, Scott, & Morency, 2012) é um tema de pesquisa recorrente entre geógrafos, urbanistas e cientistas sociais. Recentemente, os métodos e a abordagem desse campo de estudos estão sendo transformados. Isso é resultado da influência de, entre outras coisas, *Big Data* e abertura de diálogos com outras disciplinas (Schwanen, 2016). Ao mesmo tempo, o campo continua a tornar-se relevante: enquanto o atrito da distância for uma realidade, a acessibilidade continuará sendo um conceito útil para descrever a experiência urbana (Páez et al., 2012).

Considerando principalmente abordagens de *Big Data*, Letouzé e Jütting (2015) afirmam que os órgãos oficiais responsáveis pela produção de estatísticas e indicadores oficiais, incluindo instituições acadêmicas, devem estar cientes das evoluções do *Big Data*. Isso é necessário tanto para lucrar com novas ferramentas e abordagens, o que vem junto de rigor científico da validação e análise, como para encarar o mundo do *Big Data* como uma fonte inestimável de dados para o avanço da pesquisa científica. O potencial dos instrumentos de *Big Data* também é de fundamental importância para os gestores públicos e os decisores políticos (Kim, Trimi, & Chung, 2014). A utilização de diferentes ferramentas de *Big Data* numa rede de transportes multimodais pode ajudar a regular melhor a oferta de transporte privado e a fornecer melhor transporte público de acordo com as necessidades dos utilizadores (Kim et al., 2014; Lessa, Lobo, & Cardoso, 2019).

Nesse contexto, o crescimento das empresas tecnológicas globais de mobilidade, tais como Uber, Cabify e Lyft, tornou-as atores consideráveis nessa área. Por exemplo, a Uber realiza uma média diária de 15 milhões de corridas em todo o mundo. No Brasil, a empresa oferece seu serviço em 100 cidades, com uma rede de 500 mil motoristas e mais de 20 milhões de usuários.

A Uber desenvolve, comercializa e opera o aplicativo para *smartphones* que permite aos consumidores solicitar corridas de motoristas parceiros. No processo de solicitação de corridas, a ferramenta Uber fornece estimativas de tempo de espera e custo de viagem no aplicativo do usuário e no ambiente "web" por meio de interfaces de programação de aplicações (APIs) públicas.

Essas APIs geram estimativas por meio da análise do histórico de corridas na região do usuário e da curva de oferta e demanda dos carros da Uber (Cohen et al., 2016), conforme pode ser verificado na documentação disponível (https://developer.uber.com).

Wang & Mu (2018) e Hughes & Mackenzie (2016) propõem o uso dessas estimativas como uma possível medida de acessibilidade. As interferências desses novos serviços no ambiente de transporte e o fácil acesso às ferramentas que a Uber fornece já motivaram a realização de alguns estudos (Hall & Krueger, 2016; Hughes & MacKenzie, 2016; Wang & Mu, 2018; Zhou, Wang, & Li, 2017). Isso abre oportunidade para a exploração empírica da frota da Uber à luz da teoria da acessibilidade.

O objetivo deste estudo é utilizar as ferramentas de *Big Data* desenvolvidas pela Uber, uma das maiores provedoras de terceirização de viagens de São Paulo, para gerar dados para a

 CC BY © RAE | São Paulo | **V. 59** | **n. 6** | **nov-dez 2019** | 402-414 ISSN 0034-7590; eISSN 2178-938X

realização de um estudo exploratório possivelmente mensurando acessibilidade.

Como o algoritmo de preços da Uber segue o equilíbrio entre oferta e procura, a fim de influenciar o comportamento do motorista (Hall, Horton, & Knoepfle, 2019), inferimos que os tempos de espera para corridas na Uber podem refletir desequilíbrios regionais na oferta de carros. Seguindo essa linha de pensamento, exploramos a relação dos tempos de espera com outras variáveis associadas à acessibilidade em literatura similar (Hughes & MacKenzie, 2016; Lessa et al., 2019; Wang & Mu, 2018). Em especial, variáveis relacionadas a mobilidade e fatores socioeconômicos.

Os resultados encontrados são divergentes dos encontrados em estudos anteriores (Hughes & MacKenzie, 2016; Wang & Mu, 2018) em relação à relevância das variáveis no contexto de São Paulo. Tendo em mente o alerta feito por Schwanen (2016) sobre a generalização de conclusões advindas do uso de *Big Data* em diferentes contextos locais, estudos adicionais podem comparar diferentes cidades para elucidar como o contexto local pode ser mais bem considerado em abordagens semelhantes. Além disso, uma validação adicional dessa fonte de dados pode construir uma nova ferramenta capaz de facilitar a tomada de decisão dos planejadores de transportes e gestores de políticas públicas nas cidades (Kim et al., 2014; Letouzé & Jütting, 2015).

REVISÃO DA LITERATURA E QUESTÃO DE PESQUISA

Esta seção começa com a conceitualização de *Big Data* e seu posicionamento no mercado de tecnologia atual e a conceituação da ferramenta de estimativa da Uber à luz do conceito de *Big Data*. Seguem-se uma revisão do conceito de acessibilidade, suas disposições normativas e positivistas e seu contexto gerencial.

Big Data

Uma definição clássica do movimento de *Big Data* leva em consideração as características dos dados produzidos nos ambientes virtuais com presença massiva do usuário: Volume, Variedade e Velocidade, os três Vs (McAfee & Brynjolfsson, 2012).

De acordo com essa definição, os dados gerados pelas novas formas de usar tecnologia e aplicativos geram bases de dados relativamente grandes, chegando à escala de Petabytes ou Exabytes. As diversas formas de geração de dados (como postagens com fotos, comentários em redes sociais, reações a comentários de outros, vídeos e áudios etc.) são responsáveis pela geração de bases de dados heterogêneas, em contraste com as bases de dados estruturadas. A geração dessas bases de dados é quase instantânea, exigindo processamento em tempo real em alguns casos. Esses aspectos subsidiam as técnicas capazes de lidar com os dados neste período valioso.

Essa não é a única abordagem conceitual para *Big Data*. Uma visão mais sociológica tenta descrever o movimento com três "Cs": "Crumbs", Capacidades e Comunidades (Letouzé & Jütting, 2015).

"Crumbs", ou migalhas, é uma referência à natureza dos dados coletados em relação ao comportamento dos usuários nessas novas aplicações. Esses usuários deixam vestígios de suas atividades enquanto interagem, e esses traços, ou migalhas, constituem os bancos de dados a serem analisados em *Big Data*. Capacidades são as técnicas, tanto estatísticas como de programação, utilizadas para manipular esses dados e extrair informações. O terceiro conceito, Comunidades, refere-se aos padrões de comportamento dos produtores de ambientes de *Big Data* dentro de comunidades específicas cujos membros compartilham ideias com linguagem específica e métodos de validação comuns. Essas comunidades podem ser estabelecidas em ambientes abertos e colaborativos, como comunidades open source, ou em grupos mais restritivos, como grupos de tecnologia em grandes corporações com acesso a grandes bancos de dados (Letouzé & Jütting, 2015).

A API da Uber no contexto de *Big Data*

A ferramenta de estimativa da Uber usada neste artigo é, por ambas as definições, uma ferramenta de *Big Data*. Em relação aos três Vs, o enorme volume de corridas faz com que os dados estejam constantemente alimentando o algoritmo da ferramenta com interpretações de alta velocidade de dados espaciais não estruturados (Cohen et al., 2016).

Ao mesmo tempo, os dados que alimentam o algoritmo são traços de atividades dos motoristas e usuários, ou migalhas. As ferramentas que processam os dados recebidos on-line a essa alta taxa também são ferramentas incluídas no termo Capacidades, desenvolvido por ambientes de *Big Data*. Além disso, a comunidade de operadores no sistema de startups de transporte, particularmente Uber, se encaixa no conceito de Comunidades.

Além disso, a descrição de Pääkkönen e Pakkala (2015) da arquitetura de software de *Big Data* se adapta muito bem à arquitetura e tecnologia da API da Uber, conforme bem descrito em sua documentação (https://developer.uber.com).

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

Acessibilidade

Discutir o acesso dos indivíduos à mobilidade urbana na cidade é um tema importante. Há uma série de abordagens diferentes usando desde questionários estatísticos tradicionais (Metrô, 2008) até simulações computacionais de comportamento de mobilidade urbana em determinado território urbano (Krajzewicz, Erdmann, Behrisch, & Bieker, 2012), incluindo abordagens semelhantes à atual, usando dados de mídia social para traçar o comportamento de mobilidade (Noulas, Scellato, Lambiotte, Pontil, & Mascolo, 2012).

Em grandes centros urbanos, comparáveis em tamanho e importância regional a São Paulo, conhecer a mobilidade das pessoas é essencial para compreender a dinâmica urbana, pois o tamanho da cidade alimenta sua capacidade de atrair mais pessoas (Aranha, 2005). Isso tem implicações para soluções de mobilidade que têm como objetivo atender a demanda por locomoção. Um dos principais fatores para medir a capacidade de movimentação dos indivíduos dentro da cidade é o tempo usual das corridas deles.

Em São Paulo, essa variável reflete contextos socioeconômicos, indicando que o acesso à mobilidade se distribui de modo diferente em termos de espaço e contextos sociais (Morandi et al., 2016). No entanto, a mobilidade como conceito de capacidade de movimentação na cidade é limitada no sentido de não considerar as condições de locomoção.

Um conceito mais abrangente nesse sentido é o de acessibilidade (Litman, 2003; Páez et al., 2012; Stelder, 2016). É possível definir acessibilidade como o potencial de acesso a oportunidades distribuídas espacialmente (Wang & Mu, 2018). Hansen (1959) define acessibilidade como a maneira como as pessoas interagem com lugares. Para ilustrar, a quantidade de estradas medida em quilômetros em uma determinada região é uma medida comumente utilizada. Da mesma forma, na geografia, a acessibilidade é a medida de como uma pessoa participa de uma determinada atividade (Kwan, 1998; Weibull, 1980).

A acessibilidade reflete o desenvolvimento espacial, que consiste em redes de transporte e distribuição de oportunidades, materializadas em usos e ocupações do solo. A partir desse conceito, acessibilidade pode ser interpretada como uma medida temporal (tempo de acesso) (Lessa et al., 2019; Páez et al., 2012). Como exemplo prático, é possível considerar o tempo de viagem até o trabalho como uma medida comparativa para compreender o equilíbrio na ocupação profissional e as disparidades raciais, econômicas e de gênero contidas nas distribuições da área urbana (Preston & Mclafferty, 1999; Tribby & Zandbergen, 2012).

Geógrafos e cientistas sociais têm analisado criticamente a economia e a desigualdade no transporte, e suas correlações com as desigualdades socioeconômicas. Schwanen (2016) argumenta que a distribuição de transportes intensifica a polarização socioespacial sob a dinâmica do capitalismo, porque a infraestrutura de transporte é um ativo que atrai capital e investimentos, trazendo oportunidades de emprego, mais eficiência e concorrência.

Ao discutirem sobre acessibilidade, Páez et al. (2012) definem duas abordagens epistêmicas para estudar o conceito: a primeira é uma abordagem normativa definida em termos de quais parâmetros de acessibilidade devem ser considerados razoáveis ou, em outras palavras, o quanto é razoável para uma pessoa realizar uma viagem. A segunda abordagem é o positivismo, e é definida em termos de parâmetros de acessibilidade observados ou o quanto as pessoas realizam viagens. A abordagem normativa analisa as expectativas de viagem, enquanto o positivismo se baseia na experiência de viagem real. Nosso estudo levará em consideração a abordagem positivista, já que os dados se referem aos tempos reais de espera da Uber em São Paulo.

Estudos como os de Hughes e MacKenzie (2016), Lessa et al. (2019) e Wang e Mu (2018) também utilizam a abordagem positivista. De uma forma mais tradicional, Lessa et al. (2019) investigaram as relações entre os tempos de corrida coletados em questionários de origem-destino e dados de trânsito e a distribuição da infraestrutura da rede de transportes públicos. A abordagem de Hughes e MacKenzie (2016) para acessibilidade usa tempos de espera por serviços Uber e dados socioeconômicos para explorar as correlações espaciais entre a dispersão do serviço da Uber e as disparidades socioeconômicas de Seattle. Esse dado pode ser mensurado por variáveis como densidade populacional, renda média per capita e percentual de não brancos. Wang e Mu (2018) realizaram um estudo transversal, criando um modelo de regressão espacial que testa a relação do tempo de espera pelo Uber com variáveis socioeconômicas e de infraestrutura de transporte da cidade de Atlanta.

Como Hughes e MacKenzie (2016) e Wang e Mu (2018) apontam, esses estudos preliminares estão sujeitos a influências econômicas e culturais do local pesquisado. Além disso, Letouzé e Jütting (2015), Schwanen (2016) e Kwan (2016) nos alertam sobre a possível existência de viés em bancos de dados construídos a partir de ferramentas de *Big Data*. Isso é extremamente relevante para a discussão da acessibilidade e uso de bancos de dados provenientes de *Big Data* na investigação do tempo de espera pelo Uber como proxy para acessibilidade em São Paulo, pois o contexto econômico e cultural em questão é relevante para os resultados da análise.

Inspirado nesses estudos, este artigo discute as seguintes hipóteses no contexto de São Paulo:

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

H_1: O tempo de espera estimado ao solicitar uma viagem na plataforma da Uber pode ser usado como proxy para acessibilidade.

H_2: A distribuição do tempo de espera pelo Uber está associada à polarização dos indicadores socioeconômicos.

METODOLOGIA

Esta seção começa com uma descrição dos dados coletados a respeito dos tempos estimados de espera pelo Uber e dados socioeconômicos seguidos de uma análise descritiva para definir os recortes de dados para análise final (Wang & Mu, 2018). Em seguida, explicamos o processo de construção dos modelos de regressão linear múltipla (MLR) com método "stepwise". Como as variáveis significativas selecionadas por esse método foram consideradas pelo teste I de Moran como tendo alta dependência espacial, calculamos um modelo espacial autorregressivo final (SAR), conforme sugerido por Wang e Mu (2018).

Dados e variáveis

Para verificar as hipóteses propostas acima, foram coletados dois conjuntos de dados. O primeiro são os dados dos tempos de espera pelo Uber e o segundo, dados socioeconômicos espacializados de São Paulo.

Dados Uber

Os dados foram coletados a partir das APIs presentes no portal de desenvolvedores da Uber, acessíveis a partir do *site* da empresa (https://developer.uber.com), os quais abrangem os tempos de espera estimados de todos os produtos Uber durante agosto de 2018.

A cidade de São Paulo está dividida em 96 distritos, uma divisão territorial e administrativa que proporciona à administração local um certo grau de autonomia (Francisco, 2010). Seguindo a metodologia proposta por Wang e Mu (2018), os distritos foram utilizados como unidade espacial de análise. Para garantir que cada distrito tivesse pelo menos um ponto amostral aleatório, o software desenvolvido para a coleta de dados seguiu a lógica abaixo:

- Seccionar a cidade em quadrados de 1 km², totalizando 1.720 quadrados;
- amostrar aleatoriamente um ponto em cada quadrado;
- as coordenadas amostradas são usadas para consultar a API de tempos de espera pelo Uber;
- se o retorno for bem-sucedido, o resultado é armazenado;
- se um erro for sinalizado, mais três tentativas com a mesma coordenada são feitas antes de a tentativa ser registrada como um nulo.

Esse processo foi repetido a cada 30 minutos durante agosto de 2018. Mais de 2.528.400 chamadas para a API foram armazenadas, sendo pouco mais de 2.240.000 válidas.

Para processar os dados coletados, um programa foi desenvolvido em C# seguindo os conceitos de computação em nuvem para acessar as APIs da Uber. Esse programa foi alocado no serviço de computação em nuvem do Azure fornecido pela Microsoft. Sua arquitetura é apresentada na Figura 1.

Dados socioeconômicos

A escolha dos indicadores socioeconômicos foi baseada em estudos semelhantes (Hughes & MacKenzie, 2016; Wang & Mu, 2018). As unidades de análise escolhidas foram os distritos de São Paulo. Foram encontradas variáveis comparáveis para densidade populacional, densidade de emprego, taxa de minoria, renda média per capita, taxa de motorização (para carros e motocicletas), infraestrutura de transporte público e tempo médio de viagem para o trabalho. Os dados socioeconômicos em nível distrital foram coletados a partir do censo demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE, 2010), do Ministério Nacional do Trabalho e Emprego (2016, Relatório Anual de Informação Social – RAIS), do portal de indicadores dos municípios paulistas (IMP) do Seade e do portal municipal de informação georreferenciada de São Paulo (http://geosampa.prefeitura.sp.gov.br). Esses dados encontram-se resumidos no Quadro 1.

O censo demográfico decenal é um dos produtos estatísticos mais importantes do Brasil. O censo de 2010 apresenta dois conjuntos de dados: o universal, que idealmente compreende todos os brasileiros, e a amostra, na qual é feito um levantamento mais detalhado dos entrevistados (uma fração estatística das populações). A fração amostral varia entre municípios: em São Paulo, cerca de 5% dos domicílios foram incluídos na amostra. Em todo o Brasil, foram selecionados 10,7%, ou seja, 6.192.332 domicílios (IBGE, 2010).

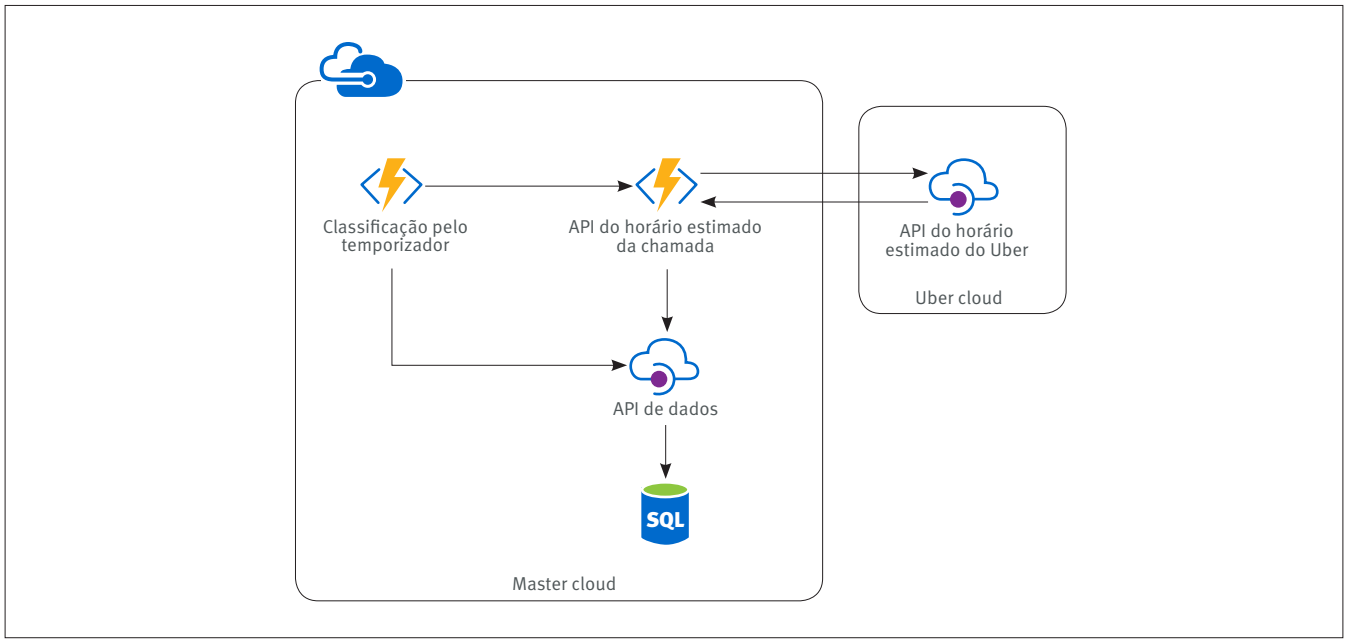
A RAIS é a ferramenta do Ministério do Trabalho e Emprego para a gestão das relações de trabalho brasileiras. Os dados são compilados a partir de declarações feitas por empresas sobre a situação de trabalho entre elas e seus funcionários. As declarações são obrigatórias para uma série de empresas nomeadas por lei (Ministério do Trabalho, 2016).

A Fundação Seade é uma instituição estatística reconhecida nacionalmente. A fundação é conhecida por sua capacidade

 CC BY © RAE | São Paulo | **V. 59** | **n. 6** | **nov-dez 2019** | 402-414 ISSN 0034-7590; eISSN 2178-938X

técnica. Também é responsável, já que faz parte do Sistema Nacional de Estatística para produção de dados, pela agregação de dados existentes no interesse do Estado de São Paulo e de seus municípios (Francisco, 2010).

Figura 1. **Diagrama da aplicação de coleta de dados da Uber**



Quadro 1. **Dados socioeconômicos de São Paulo**

Dados distritais	Fonte	Data
Área (km²)	Dados do IBGE. Coletados no portal IMP – Fundação Seade	2009
População	Dados originais do censo demográfico do IBGE reajustados anualmente pela Fundação Seade. Coletado no portal IMP – Fundação SEADE	2010/2018
Densidade da população	Calculada a partir de dados da área e população	2010/2018
Domicílios particulares permanentes	Dados originários do censo demográfico do IBGE reajustados anualmente pela Fundação Seade. Coletado no portal IMP – Fundação Seade	2010/2018
Rendimento *per capita* – censo demográfico (em reais atuais)	Dados originários do censo demográfico do IBGE. Coletado no portal IMP – Fundação Seade	2010
Empregos (comércio, serviços, indústria da transformação, construção civil)	Portal municipal de São Paulo "Infocidade". Fonte de dados original: Ministério do Trabalho e Emprego - RAIS	2010/2016
Empregadores (comércio, serviços, indústria da transformação, construção civil)	Portal municipal de São Paulo "Infocidade". Fonte de dados original: Ministério do Trabalho e Emprego - RAIS	2010/2016
% de não brancos (preto, pardos, indígenas)	Censo demográfico 2010 do IBGE	2010
Tempos de viagem	Amostra do censo demográfico do IBGE. Proporção de pessoas ponderadas em cada classe de tempo de viagem descrita no questionário, por distrito	2010
Taxa de motorização doméstica	Amostra do censo demográfico do IBGE. Taxa de motorização por distrito (carros e motocicletas) dos domicílios ponderados da amostra	2010
Paradas de ônibus	Portal Geosampa	2018
Extensão das linhas de ônibus (km)	Portal Geosampa	2018
Estações de metrô	Portal Geosampa	2018

A plataforma on-line municipal de São Paulo Geosampa fornece uma série de dados georreferenciados sobre uma série de questões, incluindo a distribuição de infraestrutura de transporte.

Todos os dados foram pesquisados ou agregados no nível do distrito (Francisco, 2010; Hughes & MacKenzie, 2016; Lessa et al., 2019; Wang & Mu, 2018). A Amostra Censo de 2010 foi utilizada para contabilizar o percentual da população de cada distrito que está em diferentes classes de tempo de viagem para trabalho/escola, e também o percentual de domicílios em cada distrito que possuem carro e motocicleta.

A distribuição de dados socioeconômicos segue um padrão claro de periferia central com pequenas variações, e está associada à distribuição de infraestrutura e políticas públicas (Francisco, 2010; Torres, Marques, Ferreira, & Bitar, 2003). As relações desse padrão com a distribuição dos tempos de espera pelo Uber podem lançar alguma luz sobre a relação de acessibilidade e composição socioespacial, como Wang e Mu (2018) investigaram.

Análise descritiva de dados

A exploração das 2.528.400 chamadas da API mostrou dois padrões: em primeiro lugar, a falta geral de estimativas para alguns dos produtos da Uber e, em segundo lugar, a ausência completa de chamadas em algumas regiões da cidade, como é mostrado na Figura 2. É possível observar a ausência no extremo sul da cidade e em uma faixa ao norte, além de alguns pontos espalhados pela cidade.

Figura 2. **Pontos de coordenada para estimativas de tempo de espera pelo Uber**



As causas identificadas foram a presença de uma área de mangue no sul sem acesso rodoviário. A faixa norte e os outros pontos sobre a cidade são regiões que a Uber classifica como zonas de risco, onde não presta serviços.

A cobertura de cada serviço e a sua disponibilidade podem ser vistas no Tabela 1. O UberX, serviço mais popular que a Uber oferece, tem a maior cobertura, com quase 100% das respostas estimadas de tempo registradas. Consequentemente, é o serviço que melhor reflete a distribuição espacial da frota da Uber. Portanto, os tempos de espera pelo UberX foram usados para medir acessibilidade.

Tabela 1. **Retornos de chamadas de estimativas de tempo de espera pelo Uber, por produto**

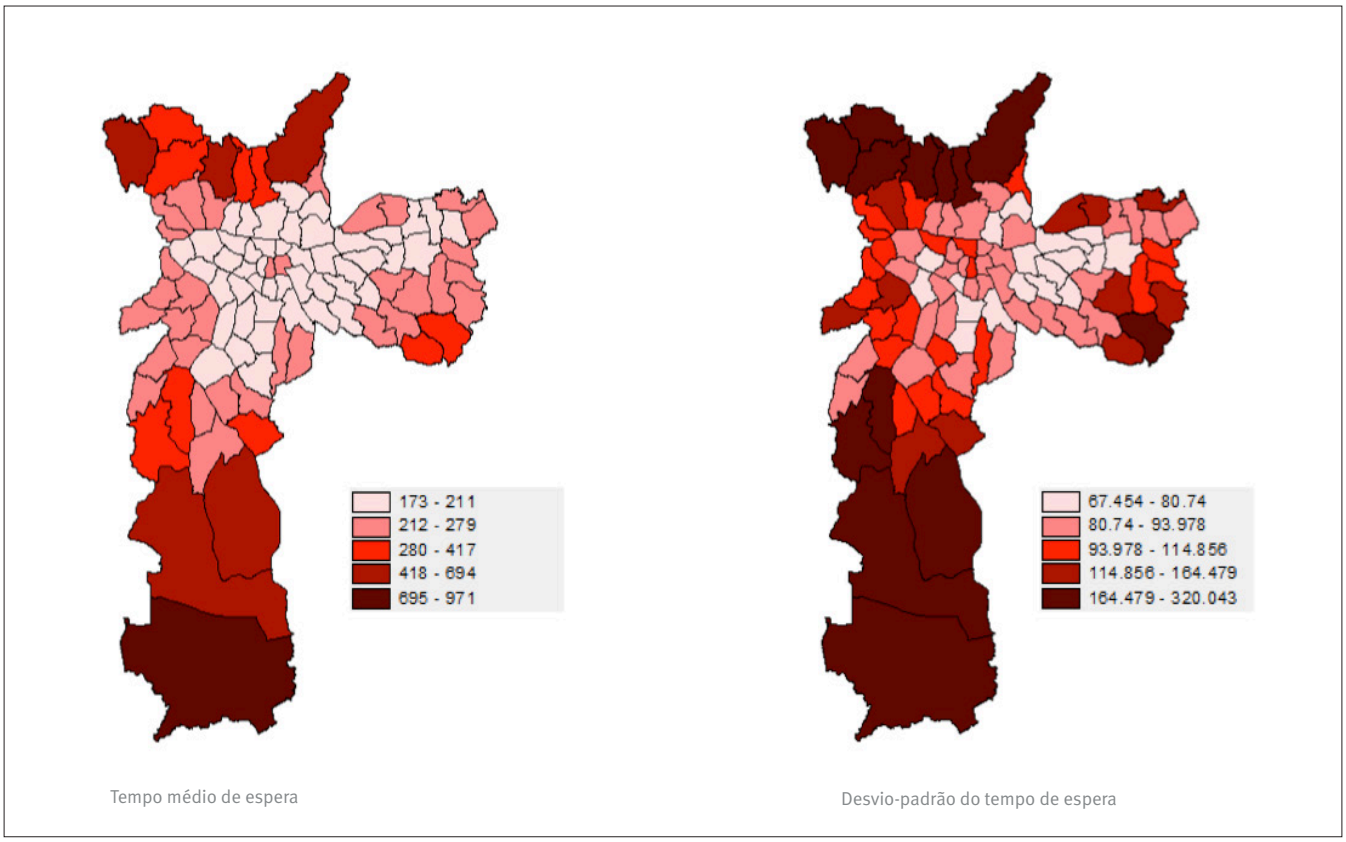
Serviço	Número de chamadas bem-sucedidas	Média do tempo de espera estimado (segundos)	Desvio-padrão do tempo de espera estimado (segundos)
Bag	1.327.710	400	187,2
Bike Rack	58.758	598	280,6
Black	987.605	491	215,5
Black Bag	549.919	506	220,6
Pool	361.166	200	87,9
Select	1.713.733	340	170,6
UberX	2.233.720	377	277,9

Dados filtrados considerando apenas o produto UberX, uma grande amplitude nas médias e os desvios-padrão entre distritos podem ser vistos na Figura 3. É razoável considerar que um distrito com uma frota maior e mais acesso apresenta menor variação nos tempos de espera. Poderíamos dizer que os distritos mais acessíveis apresentam desvios-padrão mais baixos (Wang & Mu, 2018). Por esse motivo, dois modelos MLR foram criados. No primeiro, o tempo médio de espera é a variável dependente, e, no segundo, o desvio-padrão é a variável dependente.

Regressão linear múltipla (MLR)

Com base no princípio da estatística inferencial, é possível fazer afirmações sobre as características de uma população a partir de uma amostra dela. Análise de regressão é o termo que descreve uma família de métodos que permite explorar e inferir a relação entre duas ou mais variáveis (Francisco, 2010; Hair, 2006).

Figura 3. **Comparação entre a média e o desvio-padrão dos tempos de espera estimados pelo UberX**



Tempo médio de espera

Desvio-padrão do tempo de espera

Para a construção dos modelos MLR, foram utilizados apenas os dados do UberX agregados por distritos. A variável dependente para o modelo 1 foi o tempo médio de espera, e a variável dependente para o modelo 2 foi desvio-padrão. As variáveis independentes utilizadas foram:

- área (km²)
- população
- densidade populacional
- rendimento per capita – censo demográfico (em reais)
- empregos (comércio, serviços, indústria da transformação, construção civil)
- empregadores (comércio, serviços, indústria da transformação, construção civil)
- proporção de residentes não brancos (negros, pardos e indígenas)
- tempo de viagem
- taxa de motorização de automóveis domésticos e motocicletas
- número de paradas de ônibus
- comprimento das linhas de ônibus
- quantidade de linhas de ônibus
- número de estações de metrô

O software R e suas extensões "stats" e "car" foram utilizados para o cálculo dos modelos regressivos.

Modelo espacial autorregressivo (SAR)

Francisco (2010) sugere que, antes de criar um modelo SAR, seja verificada a autocorrelação espacial da variável dependente. A literatura utiliza uma medida estabelecida por Moran. O índice de Moran é um indicador da correlação entre o valor da variável observada em uma unidade espacial de análise e os valores dessa variável na região da unidade (seus vizinhos).

Após a verificação da autocorrelação geográfica dos tempos de espera pelo Uber por meio do I de Moran, foram selecionadas as variáveis altamente significativas do modelo MLR, e o SAR foi calculado. Francisco (2010) define SAR como um modelo de regressão capaz de incorporar a matriz dos vizinhos espaciais (ou proximidade espacial) como parte das variáveis explicativas.

O software GeoDa versão 1.12 foi usado para construir o SAR.

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

RESULTADOS

A Tabela 2 resume os resultados dos modelos MLR tendo como variável dependente a média dos tempos de espera pelo UberX e o desvio-padrão dos tempos de espera. É possível notar que o modelo de tempo médio possui maior poder de explicação, pois tem um R^2 de 0,893 contra um R^2 de 0,717 do modelo do desvio-padrão.

Tabela 2. **Resultado das regressões de desvio médio e padrão para o UberX**

Variável dependente		Média do UberX			Desvio-Padrão do UberX		
		Coeficiente	Erro-padrão	valor de p	Coeficiente	Erro-padrão	valor de p
(Intercepto)		943,00	750,40	0,213	39,35	515,30	0,939
QTLINBUS2018	Quantidade de linhas de ônibus	0,13	0,19	0,514	0,12	0,13	0,353
KMLINBUS2018	Comprimento das linhas de ônibus (km)	-0,12	0,07	0,080	-0,09	0,05	0,057
QTPONTBUS2018	Número de paradas de ônibus	-0,12	0,13	0,361	0,01	0,09	0,902
QTESTMETRO2018	Quantidade de estações de metrô	0,91	5,22	0,862	-1,14	3,58	0,751
RENDP2010	Rendimento *per capita*	-0,00	0,01	0,922	0,00	0,01	0,515
ARE1	Área (km^2)	3,55	0,31	0,00	0,74	0,21	0,001
POP2018	População	0,00	0,00	0,492	-0,00	0,00	0,894
DENPOP2018	Densidade populacional	-0,00	0,00	0,007	-0,01	0,00	0,000
DOMP2018	Número de domicílios particulares permanentes	-0,00	0,00	0,646	0,00	0,00	0,546
ESTAB2016	Empregadores	-0,00	0,01	0,970	-0,00	0,00	0,791
EMP2016	Empregos	0,00	0,00	0,384	0,00	0,00	0,547
PNBRAN2010	Proporção de não brancos	300,20	77,69	0,000	202,20	53,35	0,000
TEMP2010_5 MIN	Tempo de viagem – até 5 minutos	-587,50	1.166,00	0,616	111,00	800,50	0,890
TEMP2010_30MIN	Tempo de viagem – de 6 a 30 minutos	-978,90	785,70	0,217	37,68	539,50	0,945
TEMP2010_60MIN	Tempo de viagem – de 31 a 60 minutos	-278,40	831,70	0,739	180,80	571,10	0,752
TEMP2010_120MIN	Tempo de viagem – de 61 a 120 minutos	-1,046,	805,20	0,198	-141,10	552,90	0,799
TEMP2010_121MIN	Tempo de viagem – superior a 121 minutos	- 1,443,	992,80	0,150	-215,90	681,80	0,752
TEMP2010_0MIN	Viagem não realizada	-821,30	765,00	0,286	10,86	525,30	0,984
PDOMC2010	Carro – taxa de motorização doméstica	-139,10	367,70	0,706	-164,10	252,50	0,518
PDOMM2010	Motocicleta – taxa de motorização doméstica	37,73	82,42	0,648	25,17	56,60	0,658

Um modelo de MLR *stepwise* analisando a colinearidade das variáveis independentes do modelo final foi criado. Consideramos conveniente trabalhar com fatores de inflação de variância (FIVs) inferiores a 5, conforme recomendado por Batterham, Tolfrey e George (1997). Além disso, o modelo selecionado foi a média do tempo de espera, devido ao seu maior poder de explicação. A Tabela 3 mostra o resultado do modelo. Entretanto, a Tabela 4 mostra a análise de colinearidade das variáveis independentes selecionadas para esse modelo final. Podemos observar alta significância para o comprimento das linhas de ônibus, área do distrito, densidade populacional, empregos, porcentagem de não brancos e tempos de viagem superiores a 120 minutos, com um R-quadrado de 0,879. Esses elementos corroboram a visão positivista dos estudos de Lessa et al. (2019) e Páez et al. (2012).

Tabela 3. **Resultados de regressão *stepwise* para tempos médios de espera pelo UberX**

Variável dependente		Média do UberX		
		Coeficiente	Erro-padrão	valor de p
(Intercepto)		171,00	18,17	0,000
KMLINBUS2018	Comprimento das linhas de ônibus (km)	- 0,0786	0,02	0,002
ARE1	Área (km²)	3,634	0,23	0,000
DENPOP2018	Densidade populacional	- 0,0035	0,0009	0,000
EMP2016	Empregos	0,00021	0,0001	0,097
PNBRAN2010	Proporção de não brancos	295,70	44,57	0,000
TEMP2010_121MIN	Tempo de viagem – superior a 121 minutos	- 1,239,00	422,50	0,004

Tabela 4. **Análise dos fatores de inflação de variação**

Fator de inflação de variância		Média do UberX
		Coeficiente
KMLINBUS2018	Comprimento das linhas de ônibus (Km)	1,181254
SÃO1	Área (km²)	2,139971
DENPOP2018	Densidade populacional	1,317234
EMP2016	Empregos	1,894758
PNBRAN2010	Proporção de não brancos	2,462318
TEMP2010_121MIN	Tempo de viagem – superior a 121 minutos	2,615456

Figura 4. **Tempo médio de espera pelo Uber X de acordo com o teste I de Moran**

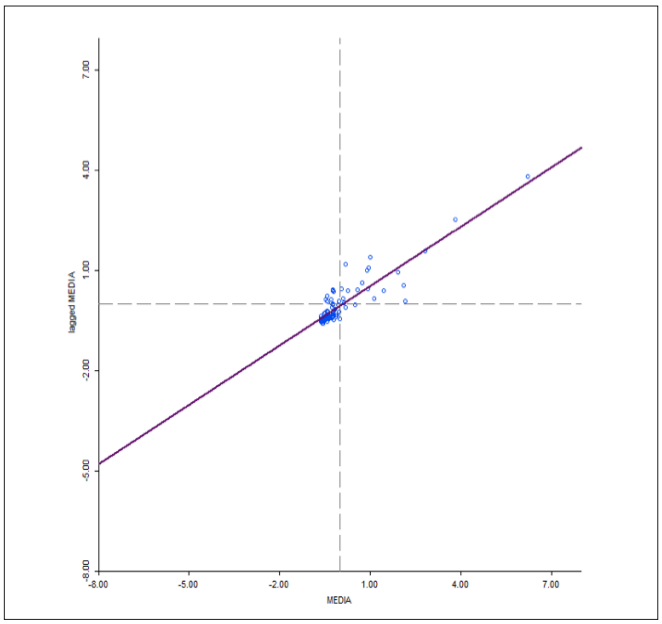
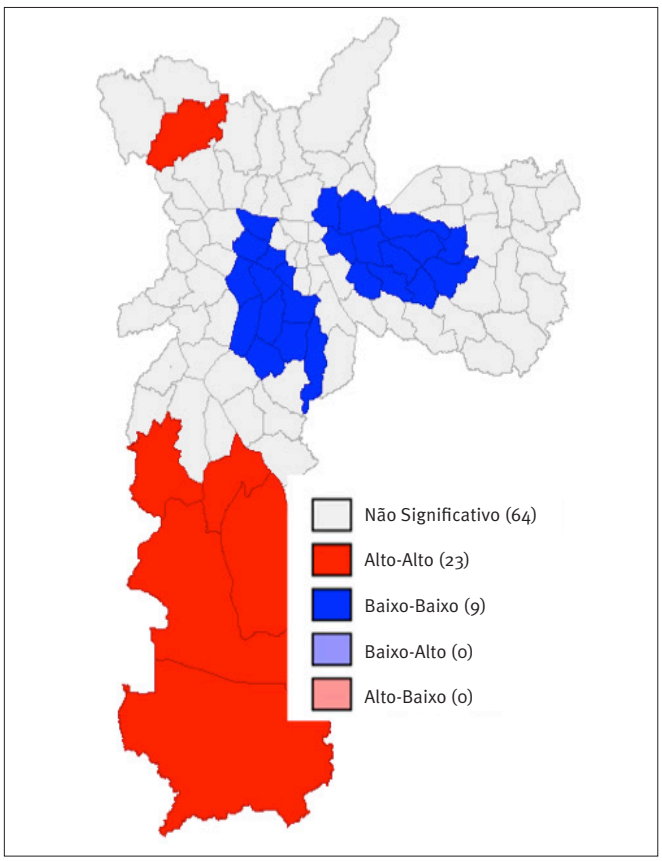


Figura 5. **Tempo médio de espera pelo UberX por *clusters* de bairros**



A Figura 4 mostra a dependência geográfica do tempo médio de espera pelo Uber X, com um I de Moran de 0,59, o que mostra a inclinação geográfica da variável dependente. Na Figura 5, observamos a distribuição da variável e sua dependência geográfica através do bairro, pois podemos notar áreas uniformes com menor tempo e baixa presença.

A partir das variáveis da Tabela 2, geramos um modelo de regressão espacial para o tempo médio de espera pelo UberX. Os resultados apresentados na Tabela 5 indicam alta significância para as variáveis comprimento das linhas de ônibus, área do distrito, densidade populacional e percentual de não brancos. Com a incorporação do fator geográfico no modelo, o R-quadrado salta para 0,89.

Tabela 5. **Resultados da regressão de atraso espacial do UberX**

Variável dependente		Média do UberX		
		Coeficiente	Erro-padrão	valor de p
W_MEDIA		0,406	0,092	0,000
CONSTANTE		114,573	23,246	0,000
KMLINBUS2018	Comprimento das linhas de ônibus (km)	- 0,073	0,022	0,001
ARE1	Área (km²)	2,455	0,263	0,001
DENPOP2018	Densidade populacional	- 0,002	0,001	0,005
PNBRAN2010	Proporção de não brancos	211,038	44,297	0,000

Observando os coeficientes do modelo SAR final, podemos dizer que o número de linhas de ônibus (-0,073) e a densidade populacional (-0,0023) influenciam negativamente o tempo de espera pelo Uber, assim como reportado por Wang e Mu (2018). A lógica é que a maior concentração de população e seu fluxo de transporte atrairiam os motoristas da Uber para essas regiões, impulsionando a oferta e proporcionando, assim, um menor tempo de espera para o serviço.

Em contrapartida, vale ressaltar que o coeficiente do percentual de não brancos (211,03) indica um aumento no tempo de espera quando a concentração de minorias é identificada na região. A possibilidade dessa relação – maiores tempos de viagem e espera *versus* distribuição de minorias – tem sido abordada de modo recorrente na literatura sobre acessibilidade (Flores & Rayle, 2017; Hughes & MacKenzie, 2016; Páez et al., 2012; Wang & Mu, 2018). Conforme sugerido por Wang e Mu (2018), as diferenças econômicas culturais entre São Paulo e Atlanta podem ser o motivo da discrepância entre os resultados obtidos. Essa ocorrência nos motiva a continuar o debate sobre as idiossincrasias regionais de acessibilidade.

Ao analisar os resultados, não podemos confirmar H_1, ou seja, não podemos afirmar que a estimativa do tempo de espera pelo serviço pode ser usada como proxy de acessibilidade. Isso se deve ao fato de que a relação com o tempo de espera pelo UberX não foi significativa para algumas variáveis.

Podemos confirmar H_2 para a distribuição da quantidade e número de linhas de ônibus, o que corrobora os resultados do estudo de Lessa et al. (2019), o qual encontrou que área do distrito, densidade populacional e o percentual de não brancos podem ser usados para medir acessibilidade. Nos estudos de Hughes e MacKenzie (2016) e Wang e Mu (2018), a relação entre o tempo de espera pelo Uber e distribuição de minorias não foi significativa.

DISCUSSÃO E CONCLUSÃO

O modelo final apresenta diferentes variáveis de estudos semelhantes. A taxa de minorias, por exemplo, apresenta uma elevada significância com um valor p de 0,00, em contraste com os casos de Seattle (Hughes & MacKenzie, 2016) e Atlanta (Wang & Mu, 2018). A relevância do contexto local é importante para explicar os achados deste estudo. Como os tempos de espera podem ter relação com a oferta e demanda consistentes de carros (Hall et al., 2019), uma explicação possível é que, como o padrão espacial socioeconômico (Torres et al., 2003) também acompanha as taxas de minorias e o status periférico, o serviço tende a ser menos acessível nessas regiões em que sua demanda poderia ser menor. No caso de São Paulo, a correlação entre taxas de minorias e renda nos encoraja a pelo menos pensar nessa possibilidade, já que a acessibilidade do Uber pode ser relativamente menor do que nos contextos de Seattle ou Atlanta. Uma futura análise usando dados das estimativas de custo da Uber pode ajudar a explorar esse padrão.

Por outro lado, algumas variáveis como densidade populacional (Hughes & MacKenzie, 2016; Lessa et al., 2019; Wang & Mu, 2018) e densidade de estradas (Wang & Mu, 2018), encontram alguma ressonância na discussão. Mesmo estando longe de ser suficiente para explicar a acessibilidade e estando cientes da natureza local do problema (Schwanen, 2017), podemos sugeri-las como variáveis que podem ajudar a descrever variáveis locais. Essa possibilidade existe porque a relação delas com a acessibilidade parece ser consistente.

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

A noção de que é necessário promover medidas de acessibilidade intuitivas e altamente replicáveis não está bem desenvolvida entre pesquisadores (Páez et al., 2012). Da mesma forma, a construção de uma medida de acessibilidade de transporte a partir de ferramentas de *Big Data*, como a API da Uber, pode contribuir para a divulgação e compreensão desse indicador para o público em geral na medida em que a teoria se aproxima de um serviço de consumo de massa. A expansão da mesma análise a outros prestadores de serviços e uma identificação mais precisa dessa participação de mercado no campo dos transportes em São Paulo também podem ajudar a identificar possíveis vieses no uso da ferramenta Uber.

A gestão pública pode fazer uso de ferramentas similares para desenvolver uma "cesta" de indicadores para diferentes modos de transporte e medições das suas inter-relações. Essas medidas podem ser utilizadas para regular melhor as atividades existentes e para adaptar a prestação de serviços públicos de mobilidade, com uma integração mais íntima entre os modos de transporte público e privado. Manter uma medida em tempo real do efeito da regulação pública em aspectos do sistema de transporte possibilitaria um melhor ajuste entre a intervenção e o interesse dos usuários (Jin et al., 2018), uma vez que indicadores bem-equilibrados podem refletir o comportamento e as necessidades dos usuários (Lessa et al., 2019; Páez et al., 2012).

No entanto, algumas limitações do estudo merecem ser destacadas, pois fornecem indicações para futuras pesquisas. É importante sublinhar a importância de uma análise mais abrangente. Este estudo limitou-se à frota da Uber. Em cidades como São Paulo, onde há pelo menos mais um ator com uma frota significativa, seria interessante replicar a metodologia com outras empresas que prestam serviços similares. Em segundo lugar, a replicação do estudo em outras cidades brasileiras parece necessária, pois os fatores econômicos e culturais que afetam as disparidades podem ser analisados em contextos mais semelhantes aos de Atlanta e Seattle, por exemplo. Em terceiro lugar, é necessário aprofundar a compreensão da ferramenta de estimativa de tempo da Uber para identificar possíveis vieses que poderiam ter suprimido alguma relação. Compará-la com outros serviços pode ser útil nesse sentido. Uma melhor compreensão desses dados pode representar um grande avanço em termos de ferramentas que possibilitem aos municípios atenderem melhor às necessidades de transporte de sua população. Além disso, vale a pena revisitar este estudo após a publicação do Censo do IBGE de 2020, pois muitas das variáveis utilizadas no estudo são projeções feitas a partir do Censo de 2010.

REFERÊNCIAS

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

FÓRUM

Enviado 24.07.2018. Aprovado 19.07.2019
Avaliado através de um processo de revisão duplo-cego. Editores Científicos: Eduardo de Rezende Francisco, José Luiz Kugler, Soong Moon Kang, Ricardo Silva e Peter Alexander Whigham
Versão original
DOI: http://dx.doi.org/10.1590/S0034-759020190607

FATORES QUE AFETAM A ADOÇÃO DE ANÁLISES DE *BIG DATA* EM EMPRESAS

Factors affecting the adoption of big data analytics in companies

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

RESUMO

Com a quantidade total de dados duplicando a cada dois anos, o baixo preço da computação e do armazenamento de dados torna a adoção de análises de Big Data (Big Data analytics - BDA) desejável para as empresas como uma ferramenta de obtenção de 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) adaptada para o contexto do BDA, adicionando duas variáveis: resistência ao uso e risco percebido. Usamos o grau de implementação dessas técnicas para dividir as empresas em usuárias e não usuárias 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 implementar BDA. Enquanto as empresas que planejam usar Big Data esperam resultados significativos, 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.

ABSTRACT

With the total quantity of data doubling every two years, the low price of computing and data storage makes the adoption of Big Data analytics (BDA) desirable for companies, as a tool to get competitive advantage. Given the availability of free software, why haven't some companies adopted these techniques? To answer this question, we extended the unified theory of technology adoption and use of technology model (UTAUT) to adapt it for the BDA context by 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.*

RESUMEN

Con la cantidad total de datos duplicándose cada dos años, el bajo precio de la informática y del almacenamiento de datos hace que la adopción del análisis Big Data (BDA) sea altamente deseable para las empresas, como una herramienta para obtener 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) para adaptarlo al 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 squares (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 y uso de tecnología, resistencia al uso, riesgo percibido.*

JUANO-PEDRO CABRERA-SANCHEZ[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, Sevilha, Espanha

INTRODUÇÃO

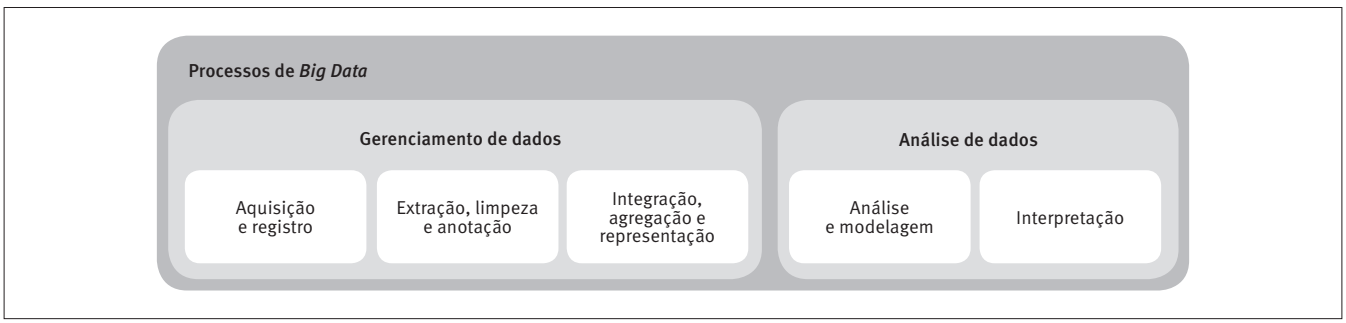
Atualmente, a sociedade gera dados sobre nossas atividades a uma taxa exponencial de crescimento. Esses dados abrangem, por exemplo, os nossos celulares e a sua localização, quaisquer transações *on-line*, a internet das coisas, redes sociais, *wearables* etc. As empresas que podem transformar esses dados sobre seus clientes em informações em tempo real ganham uma vantagem competitiva substancial (Sivarajah, Kamal, Irani, & Weerakkody, 2017). Os dados do usuário permitem que as empresas saibam quando seus clientes consomem seus produtos, os melhores momentos para promoções e como melhorar os sentimentos em relação à marca. As empresas que usam análises de *Big Data* (*Big Data Analytics* - BDA) (McAfee & Brynjolfsson, 2012) podem processar grandes quantidades de dados, quase em tempo real, e se tornar líderes de mercado.

A adoção, implementação e gestão de BDA exige que as empresas adquiram novas habilidades. Novos perfis de carreira, como cientista de dados, que combina engenharia, estatística e um profundo conhecimento de negócios, estão entre os empregos mais procurados hoje em dia. Os funcionários com essas habilidades ajudam as empresas a extrair dados gerados pelas próprias empresas e seus clientes. Isso muda a forma como as decisões são tomadas, favorecendo uma abordagem orientada a dados em oposição a uma abordagem baseada na experiência pessoal dos CEOs (McAfee & Brynjolfsson, 2012).

As empresas que utilizam BDA devem enfrentar os desafios que surgem no chamado ciclo de vida dos dados: dúvidas sobre os próprios dados, dificuldades no processamento dos dados e preocupações sobre a sua gestão (Akerkar, 2014; Zicari, 2014). As dúvidas sobre dados giram em torno de seu volume, variedade, velocidade, veracidade, volatilidade, valor e visualização. As tarefas de processamento de dados incluem técnicas relacionadas a aquisição de dados, armazenamento em bases de dados, limpeza e transformação de dados existentes, seleção correta do modelo e apresentação de resultados. Finalmente, o gerenciamento adequado de dados envolve considerações éticas, incluindo o respeito pela privacidade e segurança do usuário.

Como a tomada de decisão é cada vez mais orientada por dados, as empresas devem obter informações valiosas de modo eficiente a partir de um ambiente de dados em rápida mudança. Esse processo é detalhado por Agrawal, Bernstein e Bertino (2011), como mostra a Figura 1.

Figura 1. **Processos de *Big Data***



Como vemos, o uso de BDA pelas empresas envolve dois processos principais: gerenciamento de dados e análise de dados. Enquanto o gerenciamento de dados levanta questões de engenharia, a análise de dados está mais relacionada aos interesses dos profissionais de *marketing* e *analytics*. O BDA é o processo de obtenção de valor a partir de dados, encontrando padrões ocultos que apoiam a tomada de decisão orientada por dados.

As empresas que consideram a adoção de BDA enfrentam diversas barreiras, como falta de conhecimento, medo, resistência à mudança e limitações próprias da tecnologia (Yaqoob et al., 2016). No entanto, o BDA melhora a tomada de decisão delas, utilizando técnicas e *softwares* livres e de código aberto. Isso nos leva a dois questionamentos. Primeiro, o que afeta sua adoção? Segundo, por que muitas empresas ainda não usam o BDA? A maior parte da literatura sobre BDA concentra-se em aspectos técnicos relacionados ao seu ecossistema: desenvolvimento de aplicações, mineração de dados, análise, previsão, prescrição ou modelagem estatística (Sivarajah et al., 2017). Há pouca pesquisa sobre a adoção de BDA dentro de empresas (Brünink, 2016; Demoulin & Coussement, 2018; Huang, Liu, & Chang, 2012; Kwon, Lee, & Shin, 2014; Rahman, 2016; Verma, Bhattacharyya, & Kumar, 2018).

Este estudo, baseado no modelo de teoria unificada de aceitação e uso de tecnologia (UTAUT) (Venkatesh, Morris, Davis, & Davis, 2003), considera o impacto de duas novas variáveis,

resistência ao uso e percepção de risco, na adoção de BDA. O objetivo deste estudo é explicar a adoção e uso dessa nova tecnologia por empresas e compreender os problemas de implementação, a fim de propor recomendações aos profissionais. É por isso que fizemos uma diferenciação entre empresas usuárias e não usuárias dessa tecnologia e procuramos diferentes fatores que afetam sua aceitação e uso.

A segunda seção do artigo descreve os fundamentos teóricos do modelo proposto. A terceira seção descreve a metodologia que usamos. A quarta seção analisa os resultados obtidos a partir da aplicação do modelo em uma amostra de empresas. Por fim, apresentamos as principais conclusões teóricas e práticas, bem como as limitações da nossa pesquisa.

ANTECEDENTES TEÓRICOS E HIPÓTESES DO MODELO PROPOSTO

Análise e *marketing* usando *Big Data*

O Big Data revolucionou o *marketing* analytics e o *marketing* em geral. Forneceu novos conceitos e novas formas de fazer as coisas (Watson, 2019; Wedel & Kannan, 2016) para gerar uma vantagem competitiva. O BDA permite a inovação de serviços, o que cria valor estratégico para as empresas (Chiang, Grover, Liang, & Zhang, 2018).

Atualmente, o BDA está sendo usado em campanhas de *marketing* com o foco em aumentar o envolvimento do cliente (Liu, Shin, & Burns, 2019). A revisão da literatura mostra um número crescente de publicações sobre o uso de técnicas de *Big Data* para criar benefícios de *marketing* relacional. (Amado, Cortez, Rita, & Moro, 2018).

A gestão de *marketing* pode usar a enorme quantidade de dados disponíveis (por exemplo, nas mídias sociais) para obter *insights* valiosos de seus clientes. As empresas que exploram o *Big Data* a partir das mídias sociais ganham vantagens competitivas porque conhecem melhor os clientes (Ducange, Pecori, & Mezzina, 2018). Estudos mostram que o uso do BDA na inteligência de negócios (Sun, Sun, & Strang, 2018) e para manter a privacidade do cliente (Palmatier & Martin, 2019) cria ativos importantes no *marketing* de relacionamento.

No entanto, a literatura de adoção do BDA é relativamente escassa, e está focada no nível da indústria (Lai, Sun, & Ren, 2017; Rehman, Chang, Batool, & Wah, 2016; Wright, Robin, Stone, & Aravopoulou, 2019; Yadegaridehkordi et al., 2018). Poucos autores têm pesquisado quais fatores afetam a adoção do BDA por empresas.

Modelos de aceitação da análise de *Big Data*

A adoção de tecnologia por empresas e consumidores é fundamental para o seu sucesso. Vários modelos de adoção de tecnologia foram desenvolvidos e testados, incluindo a teoria do comportamento planejado (TPB) (Ajzen, 1991) e o modelo de adoção de tecnologia (TAM) (Davis, 1985). Mas, sem dúvida, o modelo UTAUT (Venkatesh et al., 2003) é o mais abrangente. Esse modelo integra modelos e teorias anteriores para analisar a adoção e aceitação de tecnologias.

Estudos anteriores investigando a adoção de BDA por empresas (Brünink, 2016; Demoulin & Coussement, 2018; Huang et al., 2012; Kwon et al., 2014; Rahman, 2016; Verma et al., 2018) usaram o TAM original (Davis, 1985), TAM2 (Venkatesh & Davis, 2000), TAM3 (Venkatesh & Bala, 2008) ou o modelo UTAUT sem quaisquer variáveis extras. Os modelos de aceitação foram atualizados desde a sua introdução até mesmo evoluindo para novos modelos. Como o modelo UTAUT já é um modelo maduro, nós o aprimoramos com duas novas variáveis (descobertas significativas desta pesquisa), que ajudam a explicar quando as empresas optam por adotar o BDA.

A teoria do raciocínio comportamental (Claudy, Garcia, & O'Driscoll, 2015) fornece um quadro geral em que o envolvimento do usuário é muito importante para a adoção bem-sucedida de tecnologia (Ives & Olson, 2008). Os usuários que estão predispostos a mudar têm menos resistência à adoção de uma nova tecnologia (Laumer, Maier, Eckhardt, & Weitzel, 2016). Diferentes atitudes moldam o processo de adoção de uma nova tecnologia (López, Rodríguez, & Cerveró, 2006). Esta pesquisa nos levou a buscar padrões diferentes entre empresas usuárias e não usuárias em nossa amostra.

O modelo que propomos inclui quatro variáveis independentes extraídas do modelo UTAUT. A primeira, expectativa de desempenho, é definida como o grau em que se espera que o uso da tecnologia ofereça benefícios para a empresa. A segunda, expectativa de esforço, mede a facilidade de uso esperada de uma tecnologia. A terceira, influência social, mede como os indivíduos acreditam que amigos e familiares gostariam que eles usassem uma determinada tecnologia. A quarta, condições facilitadoras, é definida como o quanto os consumidores acreditam que recursos e apoio estarão disponíveis para desenvolver um determinado comportamento. O modelo propõe uma influência direta das três primeiras variáveis sobre a intenção comportamental, enquanto as condições facilitadoras afetam a intenção comportamental e o comportamento de uso. Arenas-Gaitán, Peral-Peral e Villarejo-Ramos (2016) apontam que o diferencial desse modelo é sua capacidade de identificar quais

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

fatores são os principais determinantes da adoção. O modelo permite a inclusão de diferentes variáveis moderadoras que interferem na influência dos construtos-chave do modelo.

Adicionamos resistência ao uso e risco percebido aos construtos do UTAUT. Resistência ao uso consiste em reações negativas à mudança ou implementação de um novo sistema (Kim & Kankanhalli, 2009). O risco percebido é o potencial de perdas resultante da implementação de uma nova tecnologia ou sistema de informação (Featherman & Pavlou, 2003).

Hipóteses do modelo proposto

Propomos diversas hipóteses baseadas no modelo estendido da UTAUT para a aceitação e uso do BDA por empresas.

Expectativa de desempenho refere-se à percepção do desempenho que a tecnologia terá e é um dos construtos mais influentes em relação à intenção comportamental. Vários estudos (Brünink, 2016; Chauhan & Jaiswal, 2016; Yu, 2012) além do trabalho original (Venkatesh et al., 2003) confirmam essa relação positiva. Portanto, propomos como hipótese:

> H1: Expectativa de desempenho influencia positivamente a intenção comportamental de usar BDA.

Expectativa de esforço refere-se à expectativa de facilidade de aprendizagem e uso dessa nova tecnologia. De acordo com o modelo UTAUT, o grau de adoção do BDA depende da expectativa da facilidade ou dificuldade de usá-la. Diversos estudos confirmam essa relação (Al-Gahtani, Hubona, & Wang, 2007; Chauhan & Jaiswal, 2016; Kim, Chan, & Gupta, 2007; Lee & Song, 2013; Yu, 2012) e o efeito da expectativa de esforço na intenção comportamental. Assim, propomos como segunda hipótese do modelo:

> H2: Expectativa de esforço, ou facilidade de uso, afeta positivamente a intenção comportamental de usar BDA.

Influência social, conceito originalmente desenvolvido por Venkatesh et al. (2003) e ampliado no UTAUT2 (Venkatesh, Thong, & Xu, 2012), mede o efeito gerado pelo que os outros (amigos e familiares) pensam sobre essa tecnologia. Em um ambiente profissional, o que os gestores e colegas pensam também é muito importante (Al-Gahtani et al., 2007; Brünink, 2016; Chauhan & Jaiswal, 2016; Gupta, Huang, & Niranjan, 2010; Kim et al., 2007; Lee & Song, 2013). Portanto, propomos como hipótese:

> H3: Influência social afeta positivamente a intenção comportamental de usar BDA.

Resistência ao uso consiste na oposição ou reações negativas à implementação de uma nova tecnologia. Como Gibson (2004) conclui, a introdução de muitas tecnologias novas falhou devido à oposição dos usuários à sua implementação. Embora a literatura atual reconheça a resistência ao uso (Kim & Kankanhalli, 2009; Lapointe & Rivard, 2007), poucos estudos a integram no modelo UTAUT. No entanto, existem precedentes para usá-la para explicar a intenção comportamental (Hsieh, 2015). Norzaidi, Salwani, Chong e Rafidah (2008) encontraram relação entre a resistência ao uso e o uso do usuário, o que foi confirmado por outros estudos que não utilizaram o modelo UTAUT (Bhattacherjee & Hikmet, 2007; Poon et al., 2004). Portanto, propomos como hipótese:

> H4: Resistência ao uso afeta negativamente a intenção comportamental de usar BDA.

Risco percebido consiste no potencial de perdas na implementação de uma nova tecnologia. Além do trabalho de Featherman e Pavlou (2003), que inclui a escala de medição que utilizamos, muitos estudos consideraram risco percebido como um antecedente com impacto negativo na intenção comportamental (Kim, Ferrin, & Rao, 2008; Lee & Song, 2013; Martins, Oliveira, & Popovič, 2014). Portanto, propomos como hipótese:

> H5: Risco percebido afeta negativamente a intenção comportamental de usar BDA.

Condições facilitadoras são favoráveis quando há fácil acesso aos recursos necessários para a utilização de uma nova tecnologia e o suporte subsequente (Venkatesh et al., 2003). Em estudos posteriores utilizando UTAUT2, Venkatesh et al. (2012) descobriram que esse construto tem um efeito significativo sobre a intenção comportamental de usar uma nova tecnologia. Além disso, estudos mais recentes têm verificado efeito positivo dessa variável sobre a intenção comportamental (Duyck et al., 2010; Hung, Wang, & Chou, 2007; Wu, Tao, & Yang, 2007). Assim, propomos como hipótese:

> H6: Condições facilitadoras influenciam positivamente a intenção comportamental de usar BDA.

Tanto a TPB (Ajzen, 1991) como o UTAUT (Venkatesh et al., 2003) têm sido usados para demonstrar como as condições facilitadoras favoráveis afetam positivamente o uso de uma nova tecnologia. Vários trabalhos subsequentes (Al-Gahtani et al., 2007; Brünink, 2016; Chauhan & Jaiswal, 2016; Duyck et al., 2010; Kim et al., 2007) confirmam essa relação. Portanto, propomos como hipótese:

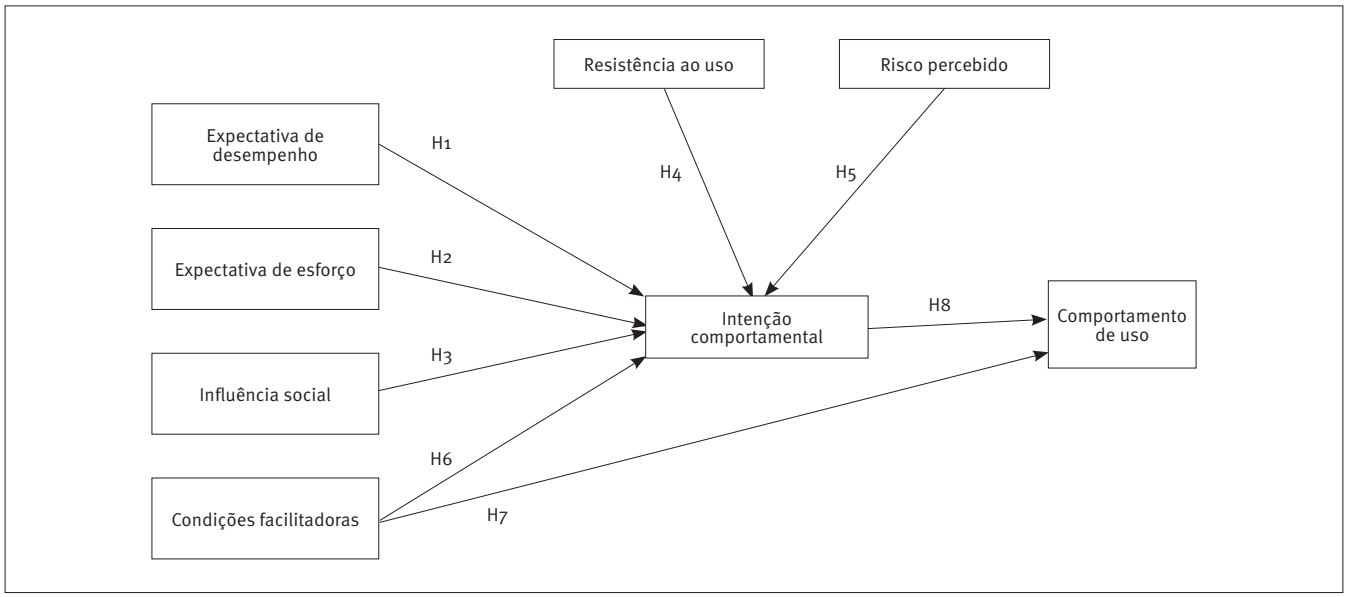
> H7: Condições facilitadoras afetam positivamente o uso de BDA.

Os principais modelos de aceitação tecnológica (TRA, TAM, UTAUT e UTAUT2) apontam para uma relação direta entre intenção comportamental e uso de tecnologias (Davis, 1985; Fishbein & Ajzen, 1975; Venkatesh et al., 2003, 2012). Essa influência tem sido demonstrada em contextos semelhantes à adoção de BDA, como *internet banking* (Martins et al., 2014), compras de voos *on-line* (Escobar-Rodríguez & Carvajal-Trujillo, 2014), sistemas eletrônicos de gerenciamento de documentos (Afonso, Gonzalez, Roldán, & Sánchez-Franco, 2012) e *Enterprise Resource Planning* (ERP) (Chauhan & Jaiswal, 2016). Portanto, propomos como hipótese:

> H8: A intenção comportamental de usar BDA afeta positivamente seu uso.

Na Figura 2, mostramos o modelo proposto de aceitação e uso de BDA com caminhos identificados em nossas hipóteses.

Figura 2. **Modelo de aceitação de *Big Data* em empresas**



METODOLOGIA DE PESQUISA

Descrição da amostra

Nossa pesquisa, cujos dados foram coletados entre setembro e outubro de 2017 via *e-mail* pessoal e telefone, contou com 199 respostas de CEOs e gerentes de empresas atuando em diferentes áreas, como Recursos Humanos, Finanças, *Marketing* e Vendas. Um pré-teste da pesquisa foi realizado com cinco gestores voluntários e vários pesquisadores especialistas, que preencheram o questionário e forneceram *feedback* sobre as questões. Na Tabela 1, classificamos as empresas dos entrevistados de acordo com suas receitas e setores.

Escalas de medição

Para os construtos UTAUT, adaptamos as escalas de Venkatesh et al. (2003) para o contexto do BDA. A resistência ao uso foi medida com a escala proposta por Bhattacherjee e Hikmet (2007), enquanto o risco percebido foi medido com a escala de Featherman e Pavlou (2003). Escalas Likert de sete pontos foram utilizadas em todos os casos.

Ferramentas estatísticas

Para estimar o modelo estrutural, utilizamos o método de mínimos quadrados parciais (PLS), (Chin & Dibbern, 2010; Hair, Sarstedt, Ringle, & Mena, 2012) com o *software* estatístico Smart PLS 3.2.3 (Ringle, Wende, & Becker, 2015). Para evitar viés de medida, ou *common method bias* (CMB), na amostra observada, seguimos as recomendações de Burton-Jones (2009). Também seguimos as recomendações de Podsakoff, MacKenzie, Lee e Podsakoff (2003), MacKenzie, Podsakoff e Podsakoff (2011); Podsakoff, MacKenzie e Podsakoff (2012), e Kock e Lynn (2012). Porque o estudo se concentra em CMB e PLS para modelos de equações estruturais. Seguindo Kock (2015), adicionamos questões não relacionadas para criar uma variável latente e consideramos esses indicadores e as demais variáveis como antecedentes. A variável CMB atua como a variável dependente de todas as outras variáveis do

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

modelo. Os fatores de inflação de variância estimados por esse método devem ser inferiores a 3,3 para confirmar que a amostra não tem CMB. Na Tabela 2, demonstramos que nossa amostra está em conformidade com esse requisito.

Tabela 1. **Empresas da amostra de acordo com a receita e setor de atividade**

	‹ 2 M€	2M€‹X‹10M€	10M€‹X‹43M€	› 43 M€	(não respondido)	- Total
Agricultura	1	3	2	1		7
Comércio e distribuição	5	4	1	10		20
Telecomunicações	6	2	4	14	1	27
Construção	2		1	4		7
Educação	2	1		2		5
Energia	1			3		4
Financeiro	1		2	8		11
Industrial	5	3	2	6		16
Outros	10	10	6	13	2	41
Saúde	3			2		5
Serviços	24	12	9	10		55
(não respondido)					1	1
Total geral	60	35	27	73	4	199

Tabela 2. *Variance inflation factor* **(VIF) de todas as variáveis para verificar CMB**

	Variável_CMB
Intenção comportamental	2,423
Expectativa de Esforço	1,631
Condições facilitadoras	2,472
Risco percebido	1,288
Expectativa de desempenho	1,994
Resistência ao uso	1,852
Influência social	1,675
Comportamento de uso	1,996

RESULTADOS

Verificamos a confiabilidade de todos os construtos. A literatura atual sugere que, para que os modelos de medição sejam considerados confiáveis e válidos, as cargas fatoriais devem exceder 0,7 (Henseler, Ringle, & Sarstedt, 2014; Roldán & Sánchez-Franco, 2012). Na Tabela 3, mostramos que cada carga foi superior a 0,7, exceto no caso do terceiro indicador de condições facilitadoras (CF3), o qual foi descartado.

Tabela 3. **Confiabilidade das escalas de medição (cargas)**

	Intenção comportamental	Expectativa de esforço	Condições facilitadoras	Risco percebido	Expectativa de desempenho	Resistência ao uso	Influência social	Comportamento de uso
IC1	0,970							
IC2	0,986							
IC3	0,984							
IC4	0,974							
EE1		0,777						
EE2		0,887						
EE3		0,898						
EE4		0,874						
EE5		0,806						
CF1			0,874					
CF2			0,892					
CF4			0,847					
RP1				0,927				
RP2				0,924				
RP3				0,877				
ED1					0,899			
ED2					0,868			
ED3					0,909			
ED4					0,918			
ED5					0,785			
ED6					0,829			
ED7					0,784			
RU1						0,931		
RU2						0,964		
RU3						0,938		
RU4						0,902		
IS1							0,741	
IS2							0,877	
IS3							0,817	
IS4							0,792	
IS5							0,724	
CU								1,000

Em seguida, analisou-se a confiabilidade do construto utilizando indicadores de confiabilidade compostos e Alpha de Cronbach. Em todos os casos, os valores dos nossos indicadores foram acima de 0,7, conforme sugerido por Nunnally (1978). A validade convergente foi confirmada pela análise da variância média extraída. Todos os valores estavam acima do limite de 0,5 proposto por Straub, Boudreau e Gefen (2004). Esses indicadores, apresentados na Tabela 4, cumprem os requisitos.

Tabela 4. **Confiabilidade composta e validade convergente**

	Alfa de Cronbach	rho_A	Confiabilidade composta	Variação Média Extraída (AVE)
Intenção comportamental	0,985	0,986	0,989	0,958
Expectativa de esforço	0,906	0,934	0,928	0,722
Condições facilitadoras	0,841	0,843	0,904	0,759
Risco percebido	0,896	0,909	0,935	0,828
Expectativa de desempenho	0,940	0,947	0,951	0,736
Resistência ao uso	0,951	0,954	0,965	0,872
Influência social	0,851	0,874	0,893	0,627
Comportamento de uso	1,000	1,000	1,000	1,000

Em seguida, avaliamos a validade discriminante do modelo de medição de duas maneiras. Primeiro, realizamos o teste de Fornell e Larcker (Barclay, Higgins, & Thompson, 1995). Em seguida, usando a razão Heterotrait-Monotrait (HTMT) (Henseler et al., 2014), asseguramos que os valores estavam abaixo de 0,9 em todos os casos. Os resultados de ambos os testes são apresentados nas Tabelas 5 e 6.

Tabela 5. **Validade discriminante (teste de Fornell-Larcker)**

	Intenção comportamental	Expectativa de esforço	Condições facilitadoras	Risco percebido	Expectativa de desempenho	Resistência ao uso	Influência social	Comportamento de uso
Intenção comportamental	**0,979**							
Expectativa de esforço	0,384	**0,850**						
Condições facilitadoras	0,628	0,587	**0,871**					
Risco percebido	-0,331	-0,185	-0,283	**0,910**				
Expectativa de desempenho	0,544	0,434	0,373	-0,195	**0,858**			
Resistência ao uso	-0,506	-0,258	-0,343	0,408	-0,566	**0,934**		
Influência social	0,497	0,459	0,483	-0,246	0,479	-0,234	**0,792**	
Comportamento de uso	0,630	0,361	0,624	-0,276	0,402	-0,400	0,449	**1,000**

Tabela 6. **Validade discriminante (Relação Heterotrait-Monotrait -HTMT)**

	Intenção comportamental	Expectativa de esforço	Condições facilitadoras	Risco percebido	Expectativa de desempenho	Resistência ao uso	Influência social	Comportamento de uso
Intenção comportamental								
Expectativa de esforço	0,380							
Condições facilitadoras	0,690	0,649						
Risco percebido	0,349	0,202	0,324					
Expectativa de desempenho	0,559	0,449	0.411	0,206				
Resistência ao uso	0,521	0,269	0,383	0,443	0,597			
Influência social	0,532	0,507	0,567	0,297	0,524	0,254		
Comportamento de uso	0,635	0,355	0,679	0,287	0,405	0,408	0,476	

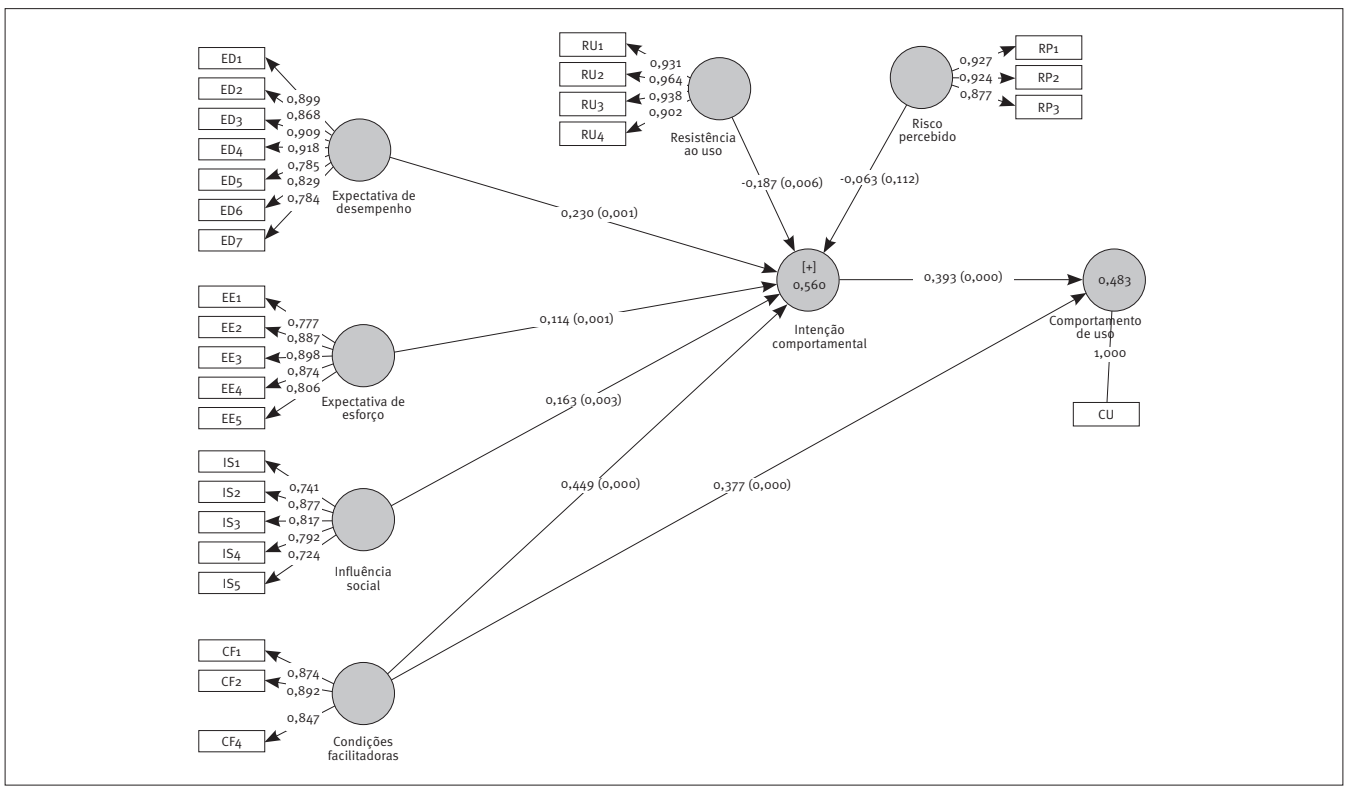
Os valores de R^2 para os construtos de segunda ordem (intenção comportamental e comportamento de uso) são mostrados na Tabela 7.

Tabela 7. **R^2 do modelo**

	R^2	R^2 ajustado
Intenção comportamental	0,560	0,546
Comportamento de uso	0,483	0,478

A Figura 3 mostra os valores para cada carga e caminho do modelo.

Figura 3. **Resultados do modelo**



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

Os valores dos coeficientes de caminho e a variância explicada das variáveis endógenas (R^2) foram analisados para avaliar o modelo estrutural. Os coeficientes de caminho indicam a intensidade da relação entre as variáveis independentes e dependentes. Utilizamos uma técnica de *bootstrapping* com 5 mil amostras para encontrar a confiabilidade dos coeficientes de caminho estimados, conforme é mostrado na Tabela 8.

Tabela 8. **Estimativas do modelo estrutural (coeficientes de caminho)**

Amostra inteira	Amostra original (O)	Valores-P
Intenção comportamental -> Comportamento de uso	**0,393 *** **	**0,000**
Expectativa de esforço -> Intenção comportamental	**-0,114 ***	**0,032**
Condições facilitadoras -> Intenção comportamental	**0,449 *** **	**0,000**
Condições facilitadoras -> Comportamento de uso	**0,377 *** **	**0,000**
Risco percebido -> Intenção comportamental	-0,063 (n. s.)	0,112
Expectativa de desempenho -> Intenção comportamental	**0,230 *** **	**0,001**
Resistência ao Uso -> Intenção comportamental	**-0,187 ** **	**0,006**
Influência Social -> Intenção comportamental	**0,163 ** **	**0,003**

***p<0,001, **p<0,01, *p<0,05. (bootstrapping com 5.000 sub-amostras e teste unicaudal).

Relações significativas com coeficientes de caminho e valor-p em negrito.

Calculou-se o indicador *Standardized Root Mean-square Residual* (SRMR) para avaliar o ajuste do modelo. O valor obtido, 0,065, foi menor que o limiar de 0,08 proposto por Henseler et al. (2014), o que sugere um bom ajuste do modelo. O modelo explica 47,85% da variação no uso e 54,6% da variação na intenção comportamental (ver Tabela 7), ambos excedendo o nível mínimo de 10% recomendado por Falk e Miller (1992).

Os resultados confirmam a maioria das hipóteses, exceto H5 (risco percebido) e H2 (expectativa de esforço). Os coeficientes das hipóteses confirmadas são significativos no nível de confiança de 1%. Embora a expectativa de esforço seja significativa no nível de confiança de 5%, encontramos uma relação negativa devido a um efeito supressor (Falk & Miller, 1992) produzido pela nova relação de condições facilitadoras na intenção comportamental, então podemos rejeitar H2. Em ordem de influência, podemos ver que o fator condições facilitadoras é o que mais contribui para a intenção comportamental e o segundo que mais contribui para o uso. O segundo fator que mais contribui para a intenção comportamental é a expectativa de desempenho, enquanto o fator que mais contribui para o uso é a intenção comportamental. Observamos também que o coeficiente de intenção comportamental de uso é significativo no nível de confiança de 0,1%. O coeficiente do efeito da resistência ao uso na intenção comportamental é significativo e negativo.

Também calculamos o Stone-Geisser Q^2 para avaliar a capacidade preditiva do modelo (Gefen, Rigdon, & Straub, 2011). Os resultados são apresentados na Tabela 9. Conclui-se que o modelo tem relevância preditiva, pois os valores de Q^2 na Tabela 9 são maiores que zero (Roldán & Sánchez-Franco, 2012).

Tabela 9. **Previsão de variáveis latentes**

	REQM	Q^2
Intenção comportamental	0,558	0,502
Comportamento de uso	0,522	0,397

Considerou-se a possibilidade de heterogeneidade na amostra. Seguindo Becker, Rai, Ringle e Völckner (2013), executamos uma segmentação de classe latente PLS-POS e também uma segmentação de classe latente FIMIX. Não encontramos diferenças em grupos com segmentação *a posteriori*.

Em seguida, tentamos várias segmentações *a priori* com critérios diferentes (por exemplo, tamanho da empresa, uso de *Big Data*, setor de atividade), não encontrando diferenças entre essas subamostras. No entanto, encontramos diferentes comportamentos nas empresas quando estas eram perguntadas sobre o nível de maturidade da implementação do BDA. Utilizamos a escala proposta por Paulk, Curtis, Chrissis e Weber (1993), que tem sido amplamente utilizada (Berg, Leinonen, Leivo, & Pihlajamaa, 2002; Khatibian, Hasan, & Jafari, 2010; Urwiler & Frolick, 2008) e tem cinco níveis: inicial, replicável, definido, gerenciado e otimizado. Foram escolhidas empresas que não tinham implementado BDA ou que estavam nos dois primeiros níveis do Segmento 1 e aquelas nos últimos três níveis do Segmento 2. Tal como indicado na Tabela 10 para o Segmento 1 e na Tabela 11 para o Segmento 2, existem diferenças significativas entre esses dois segmentos e a amostra inteira (ver Tabela 8).

Tabela 10. **Segmento 1. Estimativas do modelo estrutural (coeficientes de caminho)**

NÃO USUÁRIOS E INICIANTES Segmento 1. Tamanho: 152 empresas	Amostra original (O)	Valores-P
Intenção comportamental -> Comportamento de uso	**0,387 *****	**0.000**
Expectativa de esforço -> Intenção comportamental	-0,092 (n. s.)	0,094
Condições facilitadoras -> Intenção comportamental	**0,344 *****	**0.000**
Condições facilitadoras -> Comportamento de uso	**0,237 ****	**0,002**
Risco percebido -> Intenção comportamental	-0,073 (n. s.)	0,106
Expectativa de desempenho -> Intenção comportamental	**0,330 *****	**0.000**
Resistência ao uso -> Intenção comportamental	**-0,227 ****	**0,002**
Influência social -> Intenção comportamental	**0,189 ****	**0,002**

***p<0,001, **p<0,01, *p<0,05. (bootstrapping com 5.000 sub-amostras e teste unicaudal).
Relações significativas com coeficientes de caminho e valor-p em negrito.

Tabela 11. **Segmento 2. Estimativas do modelo estrutural (coeficientes de caminho)**

USUÁRIOS NORMAIS E *HEAVY USERS* Segmento 2. Tamanho: 47 empresas	Amostra original (O)	Valores-P
Intenção comportamental -> Comportamento de uso	**0,561 *****	**0,001**
Expectativa de esforço -> Intenção comportamental	-0,186 (n. s.)	0,084
Condições facilitadoras -> Intenção comportamental	**0,648 *****	**0,001**
Condições facilitadoras -> Comportamento de uso	0,128 (n. s.)	0,188
Risco percebido -> Intenção comportamental	-0,188 (n. s.)	0,106
Expectativa de desempenho -> Intenção comportamental	-0,214 (n. s.)	0,174
Resistência ao uso -> Intenção comportamental	-0,207 (n. s.)	0,172
Influência social -> Intenção comportamental	0,056 (n. s.)	0,354

***p<0,001, **p<0,01, *p<0,05. (bootstrapping com 5.000 sub-amostras e teste unicaudal).
Relações significativas com coeficientes de caminho e valor-p em negrito.

Realizamos um teste Anova e confirmamos as diferenças significativas entre as subamostras. Para as empresas não usuárias ou iniciantes (Segmento 1), todas as relações foram significativas, exceto os efeitos do risco percebido e da expectativa de esforço sobre a intenção comportamental. As condições facilitadoras e a expectativa de desempenho foram os fatores que mais contribuíram para a intenção comportamental, com altos níveis de significância. Os coeficientes de resistência ao uso e influência social também foram grandes e significativos. Para usuários normais e usuários pesados (Segmento 2), nenhuma das relações foi significativa, exceto o efeito das condições facilitadoras na intenção comportamental, que teve o efeito mais significativo entre todas as relações deste estudo. Curiosamente, a expectativa de desempenho não foi significativa.

DISCUSSÃO, CONCLUSÕES E LIMITAÇÕES

Nossa pesquisa amplia o modelo UTAUT para *Big Data* com uma nova variável, resistência ao uso. Com essa extensão, contribuímos para a generalização do modelo e para uma melhor compreensão da aceitação tecnológica. Nosso modelo inova, em relação à pesquisa anterior sobre BDA, ao incluir uma nova variável independente, resistência ao uso, ao modelo UTAUT e incluir o comportamento de uso de BDA como variável de desfecho. Brünink (2016) usa o modelo UTAUT sem adicionar resistência ao uso ou explicar o comportamento real de uso. Outros estudos (Demoulin & Cousssement, 2018) concentram-se no apoio do uso de aplicações de *Big Data* pela gestão, utilizando modelos como TAM, TAM2 ou TAM3 (Brown & Venkatesh, 2005; Huang et al., 2012; Verma et al., 2018). Esses modelos explicam a adoção e o uso real de BDA nas empresas, mas são mais limitados do que o modelo UTAUT.

Nossos resultados mostram que a intenção comportamental de usar BDA nas empresas é determinada por quatro fatores. Primeiro, a expectativa de desempenho, a percepção de que a implementação dessa tecnologia alcançará bons resultados, aumenta a adoção, como mostrado em estudos anteriores (Lee & Song, 2013; Yu, 2012). Segundo, a influência social tem um efeito positivo sobre a intenção de usar BDA, como demonstrado em artigos anteriores (Bozan, Parker, & Davey, 2016). Terceiro, condições facilitadoras, o apoio e o fornecimento de recursos necessários para o uso por parte da empresa aumenta tanto a intenção comportamental quanto o uso (Alharbi, 2014). Finalmente, a resistência ao uso diminui as intenções comportamentais de uso de BDA nas empresas, com um efeito mais significativo do que a influência social.

CC BY

ISSN 0034-7590; eISSN 2178-938X

Também constatamos que, embora o BDA seja percebido como sendo de difícil uso (expectativa de esforço), a influência dessa percepção sobre a intenção comportamental é pequena e contida em outras relações: condições facilitadoras na intenção comportamental (efeito supressor acima mencionado).

Também encontramos um efeito positivo das condições facilitadoras sobre o comportamento de uso da nova tecnologia com um fator semelhante para intenção comportamental. Assim, pode-se dizer que os achados são consistentes com todas as hipóteses do modelo UTAUT, exceto H5 (risco percebido). Por acharmos que resistência ao uso tem um efeito significativo, propomos acrescentá-la ao modelo original, a fim de obter melhores explicações sobre a aceitação e uso de BDA nas empresas.

Finalmente, destacamos as diferenças de comportamento entre as empresas que não utilizam essas técnicas ou estão começando a utilizá-las (Segmento 1) e as empresas que já as utilizam há muito tempo (Segmento 2). Para as empresas iniciantes ou não usuárias, a expectativa de desempenho, a influência social e as condições facilitadoras têm efeito significativo e positivo na intenção comportamental e comportamento de uso, enquanto a resistência ao uso têm efeito significativo e negativo em ambas variáveis. Entre as empresas usuárias normais e usuárias pesadas, apenas condições facilitadoras têm efeito sobre a intenção comportamental, enquanto o resto das relações não é significativo. Isso pode sugerir que os usuários estabelecidos sabem o que podem alcançar com essas técnicas, então a única coisa com que eles se importam é ter boas condições facilitadoras, enquanto os iniciantes ainda não conhecem todas as capacidades dessa tecnologia, então eles consideram mais questões.

Quanto às implicações profissionais, os resultados sugerem que os executivos inferem que cada tecnologia tem sua própria curva de aprendizagem, e essa questão não afeta sua adoção sempre que grandes resultados são esperados, como no caso do *Big Data* (Cabrera-Sánchez & Villarejo-Ramos, 2018). Da mesma forma, se a empresa tem uma infraestrutura adequada, não tem nada a perder testando a tecnologia. De qualquer forma, para superar a resistência ao uso de BDA, ainda são necessárias informações claras sobre seus benefícios. Portanto, recomendamos duas etapas para gerentes. Na primeira etapa, eles devem ser informados de que a maioria dos *softwares* associados a essas técnicas é gratuito e que, se eles já têm recursos de *hardware*, devem testá-los. Na segunda etapa, deve haver comunicação com os gestores sobre os benefícios do uso de *Big Data*, inclusive com exemplos de empresas das mesmas áreas que o utilizam. Essa segunda ação é muito importante para as empresas que estão atualmente usando *Big Data*, porque podemos inferir que elas não estão aproveitando todo o potencial da tecnologia. Elas possuem uma certa insatisfação com a tecnologia, e têm uma expectativa de desempenho muito baixa, quando o que deveria acontecer seria o oposto. Portanto, devemos informá-las sobre a tecnologia e como ela pode ser usada para gerar lucros em cada setor.

O uso de BDA nas empresas pode representar um avanço muito importante na gestão da informação, melhorando as relações com os clientes. Como é mais do que uma ferramenta de gerenciamento de relacionamento com o cliente, o BDA fornece informações relevantes às empresas e aumenta o conhecimento sobre os clientes, melhorando seu engajamento.

Embora o modelo UTAUT já tenha sido bastante testado e seja maduro, incluímos duas variáveis para estendê-lo. No entanto, muitas outras variáveis podem ser relevantes. Para essa tecnologia, construtos do modelo original, como a expectativa de desempenho, têm menor influência sobre a intenção comportamental do que um dos construtos que adicionamos, resistência ao uso. Como podem existir outras variáveis com efeitos significativos, o modelo UTAUT deve continuar evoluindo para fornecer melhores explicações para a aceitação de novas tecnologias. Pesquisas futuras sobre *Big Data* devem procurar identificar essas variáveis. Além disso, parece necessário explorar novas variáveis moderadoras com o objetivo de analisar possíveis efeitos não considerados anteriormente.

Finalmente, tamanhos amostrais maiores nos permitirão estabelecer diferenças de comportamento entre grupos de empresas, o que podemos analisar por meio de uma técnica de segmentação *a posteriori*, como segmentação orientada *a posteriori* – Mínimos Quadrados Parciais (POS-PLS). Então, se obtivermos uma amostra maior de empresas que estão usando (ou pretendem usar) *Big Data*, teremos um modelo com melhor desempenho e resultados mais informativos.

REFERÊNCIAS

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.

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

PERSPECTIVAS

Artigo convidado
Versão traduzida
DOI: http://dx.doi.org/10.1590/S0034-759020190608

BIG DATA E DISRUPÇÕES NOS MODELOS DE NEGÓCIOS

O *HYPE* DAS TECNOLOGIAS EMERGENTES

Muitos são os desafios que impedem o aproveitamento dos benefícios das tecnologias emergentes mais recentes. Se fosse simples, todas as empresas o fariam, e a vantagem competitiva desapareceria facilmente. Na verdade, é extremamente difícil e desafiador para as empresas criarem valor com tecnologias emergentes. Todos os anos, a Gartner, empresa de consultoria respeitada, analisa as tecnologias mais recentes e emergentes e utiliza o seu conceito de ciclo de *hype* para explicar o caminho que as tecnologias tomam. O ciclo consiste em cinco fases: (i) a fase de ativação tecnológica: a invenção de novas tecnologias, que ocorre em um laboratório de pesquisa, geralmente em uma universidade (hoje em dia, a maioria das empresas terceiriza suas pesquisas mais importantes para as universidades); ii) a fase de auge das expectativas inflacionadas: a tecnologia é discutida pelas empresas em conferências e na imprensa. Fala-se muito da nova tecnologia, mas ninguém a utilizou ainda. São lançados projetos de P&D; iii) a fase de desilusão: verifica-se que a tecnologia não é tão útil quanto se pensava; iv) a fase de declive de esclarecimento: é explorada a valiosa fusão entre negócios e tecnologia; (v) a fase de patamar da produtividade: fica claro como as empresas podem usar a tecnologia para criar valor. No ciclo de *hype* das tecnologias emergentes da Gartner (2019), tecnologias como biorrobôs, realidade aumentada na nuvem, *web* descentralizada, aprendizado de máquina adaptativo, impressão 3D em nanoescala e 5G são revisadas e posicionadas nas primeiras fases. A Gartner espera que essas tecnologias atinjam o patamar da produtividade dentro de 5 a 10 anos.

Duas observações importantes devem ser feitas. Em primeiro lugar, as tecnologias mais novas gerarão enormes quantidades de dados, tanto em um nível molecular quanto no nível de objetos individuais e seres humanos, e dispositivos (coisas) e seres humanos interligados incorporados em complexas redes comerciais, logísticas e financeiras. Em segundo lugar, para as empresas, seria sensato abordar essas tecnologias emergentes a partir de uma perspectiva de "detecção – armazenamento de dados – análise e resposta – aprendizagem". Do ponto de vista empresarial, o ciclo detectar-guardar-analisar-responder-e-aprender precisa estar conectado ao ciclo de *hype* das tecnologias emergentes. Completar o ciclo de "detecção-armazenamento de dados-análise e resposta-aprendizagem" é um fator crítico de sucesso para as empresas. No entanto, como de costume, há complicações; neste caso, duas. A primeira é que mais dados implicam mais ruído. É muito desafiador distinguir sinal de ruído (Silver, 2015). A segunda complicação é que os dados são gerados não apenas dentro das aplicações da empresa, mas também em aplicações executadas em conjunto com vários parceiros de negócios nos chamados ecossistemas de negócios; Porter e Heppelmann (2014) chamariam essas aplicações de negócios de "sistema de sistemas". Portanto, a propriedade e a legalidade dos dados tornam-se um fator crucial que pode impedir que as tecnologias atinjam o patamar da produtividade.

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

[1]Erasmus University, Rotterdam School of Management, Roterdão, Países Baixos

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

COMO GANHAR A CORRIDA POR DADOS?

Muitas empresas estão cientes do valor potencial das tecnologias emergentes e da transformação digital que precisam conduzir. No entanto, tal como referido anteriormente, não é fácil transformar empresas em empresas digitais. O MIT Center for Information Systems Research (CISR) tem uma longa tradição de pesquisa na análise e explicação dos chamados fatores críticos de sucesso, um termo cunhado na década de 1980 por Jack Rockart, o primeiro diretor de pesquisa do CISR. Suas últimas pesquisas mostram que as capacidades organizacionais impulsionam o sucesso mais do que a tecnologia. No seu livro *Leading digital: Turning technology into business transformation*, Westerman, Bonnet e McAfee (2014) analisaram o desempenho de 400 empresas em todo o mundo. Eles distinguiram quatro *clusters* (iniciantes, fashionistas, conservadores e mestres digitais) que se distinguem pelas capacidades que incorporam. Liderança digital refere-se ao desenvolvimento de capacidades que permitirão a concepção e implementação de aplicações de negócios úteis (O QUE) e aquelas que permitirão a transformação (COMO). O primeiro grupo, os recursos digitais, permite a criação de uma experiência de cliente atrativa, explorando o poder das operações principais e reinventando modelos de negócios. O segundo grupo, capacidades de liderança, refere-se à concepção de uma visão digital, envolvendo a organização em ampla escala, guiando a transformação e construindo capacidades de liderança tecnológica. Os recursos digitais e de liderança são condições necessárias para atingir níveis mais altos de desempenho da empresa. As empresas que incorporam essas capacidades têm um desempenho excepcionalmente melhor, tanto em termos de geração de receita quanto de lucratividade e avaliação de mercado. Os programas de transformação digital são cruciais para a sobrevivência das empresas. Veja-se, por exemplo, o Royal Philips, conglomerado estabelecido, centenário e com resultados decrescentes. Com um novo CEO em 2011, a empresa estabeleceu um programa de transformação digital, chamado *Accelerate*, que criou uma base para que a empresa se tornasse digitalmente habilitada, deixando de oferecer orientação de produtos, e para oferecer serviços de *software* (Mocker, Ross, & Heck, 2014). Um componente crucial da transformação foi eliminar a complexidade não recompensada do processo e simplificar processos, papéis e sistemas em todo o mundo. Em contrapartida, a transformação estimulou a recompensa da complexidade do produto, como soluções integradas de saúde com componentes avançados de varredura, análise de decisão e inteligência artificial. Recentemente, Peter Weill e Stefanie Woerner (2018) publicaram a última pesquisa do CISR em seu livro *What's your digital business model? Six questions to help you build the next-generation enterprise*. Eles enfatizam que a transformação digital não se trata de tecnologia, mas sim de mudança. Eles distinguem quatro modelos de negócio baseados em dimensões de *design* de negócio (cadeia de valor *versus* ecossistema) e conhecimento do cliente final (parcial *versus* completo), a saber: fornecedor, *omnichannel*, produtor modular e impulsionador do ecossistema. Os impulsionadores de ecossistema superam todos os outros modelos de negócios.

QUAL É O VALOR DO NEGÓCIO E PARA QUEM?

Existem três motivos fundamentais para as empresas criarem valor empresarial:

- Para criar negócios mais eficientes – os recursos são usados de uma maneira melhor.
- Para criar negócios que sejam mais eficazes – os clientes finais são servidos de uma maneira melhor.
- Para criar novos modelos de negócio – novos modelos de negócio geram valor sustentável.

Nessas três estratégias de criação de valor, o uso de *Big Data* e análises avançadas desempenha um papel crucial. Primeiro, veja-se, por exemplo, a eficiência do varejista na entrega da última milha de pacotes que os clientes compram *on-line*. Albert Heijn, um grande varejista na Europa, sentiu a necessidade de redesenhar o seu processo de entrega de última milha. Surgiu com uma abordagem inteligente que combinava os dados das chamadas *on-line* recebidas para entrega com o preço das faixas horárias para entrega e o agrupamento eficiente de entregas em caminhões que serviam a mesma área (Agatz, Campbell, Fleischmann, & Savelsbergh, 2011). O resultado foi uma maior utilização da capacidade dos caminhões e um número menor de quilômetros, reduzindo a pegada de carbono das entregas de última milha nas cidades.

Em segundo lugar, a Royal FloraHolland, líder mundial no cultivo de flores e plantas, está se transformando para se tornar um impulsionador de ecossistema mais eficaz no mundo da floricultura. Uma maneira é redesenhar os mercados de leilões de flores e aumentar o número de leiloeiros com capacidades analíticas avançadas. Os seus clientes, neste caso, atacadistas e varejistas profissionais, obterão os produtos certos pelo preço certo no momento certo. Devido à perecibilidade das flores, o tempo é uma fonte crucial de valor. Com base na análise de milhões de transações, eles criaram um sistema avançado de suporte à decisão que permite aos leiloeiros tomar melhores decisões e equilibrar a receita com o tempo de resposta (Lu, Gupta, Ketter, & Heck, 2019).

Em terceiro lugar, a estratégia mais desafiadora é criar negócios. A Car2Go, empresa de aluguel e partilha de automóveis,

 CC BY © RAE | São Paulo | **V. 59** | **n. 6** | **nov-dez 2019** | 430-432 ISSN 0034-7590; eISSN 2178-938X

opera em nível internacional e aluga automóveis a clientes. Em cidades como Nova Iorque, Amsterdã e Stuttgart, é possível pegar um carro na rua e deixá-lo em qualquer lugar. Um projeto-piloto com carros elétricos em San Diego, Amsterdã e Stuttgart mostrou o potencial de um novo modelo de negócio usando carros elétricos. Com a ajuda de algoritmos avançados e diferentes conjuntos de dados, como clima, preços locais de eletricidade em curto prazo e demanda de aluguel de carros, a frota de carros elétricos pode simular uma usina de energia virtual (Kahlen, Ketter, & Dalen, 2018). As baterias combinadas desses carros elétricos podem armazenar a eletricidade gerada pelo sol e pelo vento, que são fornecedores de energia renovável. Em dias nublados e/ou fechados, a eletricidade armazenada nas baterias elétricas de automóveis pode ser vendida ao público em geral por meio da rede inteligente. Assim, a Car2Go poderá mudar o seu modelo de negócio, de uma empresa de "aluguel de automóveis" para uma empresa de "aluguel de automóveis e armazenamento e venda de eletricidade".

QUAIS SÃO OS DESAFIOS TRANSFORMACIONAIS?

A transformação digital impõe muitos desafios, e, hoje em dia, a maioria das empresas enfrenta esses desafios. O primeiro é que a maioria das empresas não tem uma plataforma digitalizada comum. Elas têm uma infinidade de aplicações e sistemas que foram construídos com tecnologia e *software* relativamente desatualizados, como as aplicações de planejamento de recursos empresariais. A transformação para acesso em tempo real a enormes conjuntos de dados com recursos avançados de aprendizagem humana e de máquina levará algum tempo e esforço.

O segundo desafio é como os funcionários são capazes de trabalhar com os novos recursos avançados de inteligência de máquina. Os funcionários estão envolvidos na transformação ou se sentem ameaçados, acreditando que os robôs assumirão suas funções? Na maioria das organizações, os usuários de tecnologia não podem ou não se envolverão na transformação.

O terceiro desafio é a falta de liderança digital por parte da diretoria da empresa. A maior parte dos membros do conselho não tem conhecimento digital e, portanto, tem dificuldade de criar uma visão e missão para a empresa que seja sustentável em um mundo digitalizado. A falta de compromisso do conselho com a transformação é geralmente um sinal para o resto da empresa desacelerar o ritmo da mudança. O quarto desafio diz respeito à ética dos dados, à privacidade e à responsabilidade pelos algoritmos. As empresas adquirem dados detalhados sobre as preferências dos seus clientes. Os magnatas da tecnologia digital, como Alibaba, Amazon, Apple, Facebook, Google, Microsoft, Samsung e Tencent, reúnem dados muito detalhados sobre cada aspecto do ciclo de vida de seus clientes. Quando combinadas com ferramentas avançadas de autoaprendizagem, essas bases de dados criam poder de dados e estimulam uma discussão importante sobre as regras e responsabilidades das empresas. Quem é responsável quando as empresas criam algoritmos que discriminam funcionários ou clientes? A quem pertencem os dados gerados pelos clientes? Em maio de 2018, a União Europeia implementou um novo regulamento, denominado Regulamento Geral sobre a Proteção de Dados (*General Data Protection Regulation*), que proporciona proteção de dados e privacidade a todos os cidadãos da União Europeia e do Espaço Econômico Europeu, incluindo a transferência de dados pessoais fora do espaço europeu. Já nos outros lugares do mundo, a maior parte dos cidadãos não está protegida.

REFERÊNCIAS

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., & Dalen, J. van. (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., & Heck, E. van. (2019). Dynamic decision making in sequential business-to-business auctions: A structural econometric approach. Management Science, 65(8), 3449-3947. doi:10.1287/mnsc.2018.3118

Mocker, M., Ross, J. W., & Heck, E. van. (2014, February). Transforming Royal Philips: Seeking local relevance while leveraging global scale. *Working Paper No. 394*. MIT Sloan CISR. 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, USA: 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

PERSPECTIVAS

Artigo convidado
Versão traduzida
DOI: http://dx.doi.org/10.1590/S0034-759020190609

PLUS ÇA CHANGE, PLUS C'EST LA MÊME CHOSE

Em 1572, a análise de um único ponto de dados, a Supernova de Tycho, revelou que a abóboda celeste é inconstante, ao contrário do paradigma aceito na época (Wootton, 2015). Menos de 40 anos depois, em 1610, Galileu Galilei publicou suas sensacionais descobertas em *Sidereus Nuncius*, um curto tratado que demonstrou a existência de estrelas não vistas a olho nu e a natureza da Via Láctea (Galilei, 1610). Desde então, análises de dados têm sido fundamentais em pesquisas científicas, e há vários exemplos de seu uso na solução de problemas importantes e difíceis. Aos 24 anos, Carl Friedrich Gauss (1809) usou mínimos quadrados para prever corretamente a posição do planeta anão Ceres em 1801, ao aparecer por trás do brilho do Sol. Uma análise espacial simples identificou a fonte do surto de cólera em Broad Street, Londres, em 1854 (Snow, 1855). Entre 1856 e 1863, uma estimativa criteriosa de frequências permitiu a descoberta das regras básicas da hereditariedade de características físicas em plantas por Gregor Mendel (1866). No final da década de 1940, um estudo abrangente e retrospectivo liderado por Richard Doll e A. Bradford Hill (1950) demonstrou a forte ligação entre tabagismo e câncer de pulmão.

O constante desenvolvimento e refinamento de novos métodos estatísticos após 1880 por Galton, Student, Fisher e outros pesquisadores originou diversas aplicações de métodos de análise de dados na indústria e nos negócios. Ideias e métodos desenvolvidos e popularizados por Shewhart, Deming e outros intelectuais tornaram o controle estatístico de qualidade parte integrante do processo de fabricação industrial. Esse processo incorporou o uso de modernos modelos experimentais após a Segunda Guerra Mundial. O poder de computação de baixo custo, a coleta automatizada de dados e o desenvolvimento de *softwares* simples e versáteis de análise de dados, especialmente o R, que é abrangente e gratuito, expandiram significativamente as aplicações estatísticas. Consequentemente, nasceu a era do *Big Data* (grande volume de dados). Assumindo que grande parte do que é publicado é verdadeiro, o *Big Data* resolverá problemas críticos em áreas tão diferentes quanto diagnósticos médicos, avaliação de crédito, previsão do tempo e reconhecimento facial. Teremos produtos e serviços muito melhores, e uma compreensão muito mais profunda de processos físicos e culturais, como resultado da análise de conjuntos de dados cada vez maiores e uso de métodos robustos em computadores modernos.

Embora fazer previsões seja difícil, tenho certeza de que os principais problemas que enfrentaremos ao aplicar métodos de análise estatística a problemas de negócios na era do *Big Data* serão os mesmos problemas com os quais lidamos há décadas. Os dados analisados continuarão sendo informações, ideias e conclusões. Os dados são usados para contar uma história e analisados com métodos analíticos. Os três grandes desafios da análise de dados, em ordem de importância, são a sobrevalorização da significância estatística, a falta de reprodutibilidade e deixar de fornecer respostas exatas a perguntas erradas.

A significância estatística é vista por muitos como o padrão-ouro. Em termos simples, significância estatística é a probabilidade de que um conjunto de observações seja o resultado de

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

[1]Columbia University, School of International and Public Affairs, Nova York, NY, Estados Unidos da América

 CC BY

flutuação aleatória. A obtenção de um valor estatístico pequeno, geralmente menor que 5%, indica a presença de fatores que explicam o comportamento não aleatório. O cálculo do valor-P (a medida usual de significância) envolve um processo complexo de desenho de estudo, seleção de modelos estatísticos, e coleta e análise de dados. O nível de significância postulado é preciso apenas se todas as condições estatísticas forem satisfeitas e o modelo estatístico escolhido fornecer uma interpretação adequada dos dados. No entanto, esse raramente é o caso. As situações são desafiadoras no contexto comercial em que os modelos usados podem ser muito complexos. Esses modelos complexos são tipicamente modelos de regressão com diversas variáveis explicativas em que as relações com a resposta são consideradas lineares, e o cálculo da significância estimada dos coeficientes é frequentemente irrelevante. Coleta de dados e delineamento experimental inadequados aumentam a complexidade. O problema é tão sério que um grande movimento estabelecido para reduzir ou mesmo eliminar os valores-P reuniu muito apoio entre os estatísticos (McShane, Gal, Gelman, Robert e Tackett, 2019).

O segundo grande problema com a análise de dados é que muitos dados, talvez a maioria, não são reprodutíveis (Ioannidis, 2005). A causa mais comum é uma ou mais falhas no planejamento de um experimento ou pesquisa. Além disso, falhas na coleta de dados, uso de métodos inadequados e plágios são comuns e invalidam estudos. Em pesquisas científicas, os estudos podem ser refeitos, e algum grau de autocorreção pode ser alcançado. Contudo, em ambientes industriais, onde as decisões de negócios são frequentemente orientadas por dados urgentes, a replicação de pesquisas é mais difícil. Um estudo de mercado ou viabilidade mal realizado pode ter consequências onerosas. Um caso típico foi a plataforma Newton da Apple.

O terceiro problema pode ser o mais sério e o mais difícil de resolver. John Tukey (1962) costumava dizer que "uma resposta aproximada à pergunta certa vale muito mais do que uma resposta precisa à pergunta errada." A história clássica contada exaustivamente por estatísticos no mundo inteiro (ou "cientistas de dados", no jargão moderno), mas com mais frequência nos corredores do Departamento de Estatística da Universidade de Columbia, é sobre o reforço de aviões britânicos usados nos bombardeios da Alemanha no final da Segunda Guerra Mundial. Muitos aviões foram danificados pela artilharia antiaérea alemã e, por esse motivo, foram blindados. A preocupação mais importante a ser abordada pelos envolvidos nesse projeto foi quais partes deveriam ser blindadas. Os aviões que retornaram com avarias foram cuidadosamente examinados, e foi decidido blindar as partes que haviam sofrido mais danos pela artilharia. Curiosamente, teria sido um erro fatal se essa decisão tivesse sido tomada. A variável correta que deveria ter sido revisada não estava relacionada aos aviões que foram examinados, mas aos aviões que haviam sido destruídos e não retornaram. Felizmente, Abraham Wald inteligentemente sugeriu que as partes não afetadas nos aviões que retornaram deveriam ter sido reforçadas.

Conjuntos de dados, mesmo aqueles com *terabytes*, são apenas a matéria-prima do conhecimento. Hoje, quase tudo pode ser monitorado e medido, mas o principal desafio continua sendo a capacidade de usar e analisar conjuntos de dados e compreendê-los para fornecer interpretações confiáveis.

NOTA DA REDAÇÃO

Tradução do título: Quanto mais as coisas mudam, mais elas permanecem as mesmas

REFERÊNCIAS

Doll, R., & Hill, A. B. (1950) Smoking and carcinoma of the lung. British Medical J*ournal*, *2*(4682), 739-748. doi:10.1136/bmj.2.4682.739

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

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

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 *m*ode 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.

PENSATA

Submetido 13.03.2019. Aprovado 19.08.2019
Este artigo tem autoria de um membro do Corpo Editorial da *RAE*, foi avaliado em *double blind review*, com isenção e independência. Editor Científico: Marcelo Bispo

Versão original

DOI: http://dx.doi.org/10.1590/S0034-759020190610

CRIMES CORPORATIVOS: O ESPECTRO DO GENOCÍDIO RONDA O MUNDO

A exploração de trabalhadores, comunidades e recursos sempre ocorreu, embora não seja um tema central na história do *management*, cuja narrativa principal é a urgência pela busca da eficiência e de soluções mais acertadas. No entanto, crimes, má conduta, comportamento antiético e irresponsabilidade social corporativa têm se tornado cada vez mais comuns, ensejando reflexões sobre a relevância de abordar esses assuntos no campo acadêmico e na prática da gestão. Neste ensaio, partimos da premissa de que os crimes corporativos devem ser compreendidos pela sua natureza multidisciplinar, e, especificamente, concentramos nossos esforços para discutir questões relacionadas à gestão das organizações para argumentar que os crimes corporativos fazem parte das operações da empresa para sustentar o capitalismo contemporâneo. Iniciamos abordando o poder das corporações como a principal força do capitalismo contemporâneo, na sua forma de extrema concentração da riqueza e da propriedade das corporações. Em seguida, discutimos a seriedade dos crimes corporativos e como eles se assemelham com genocídios. Finalizamos com nossas considerações sobre o porquê de as organizações tornarem-se criminosas e fazemos uma preleção sobre uma via para que os crimes corporativos sejam evitados.

CORPORAÇÕES

A Revolução Industrial deu um novo contorno ao funcionamento das empresas, sobretudo com o surgimento de novas formas de organização de negócios, como a corporação moderna, cujo traço distintivo é a separação entre propriedade, a qual é pulverizada entre muitos acionistas, e o controle exercido por diretores que, no máximo, possuiriam uma fração pequena do capital da empresa (Berle & Means, 1932).

O surgimento das corporações no século XIX mudou os mecanismos de propriedade de empresas, inclusive no âmbito legal, e essas, rapidamente, entre 1840 e 1860, tornaram-se o modelo preferido pelos capitalistas para a organização de negócios. Nos anos 1870, as principais corporações, na maioria das indústrias, orientavam-se para reduzir a competição e aumentar seus lucros por meio de uma combinação horizontal que lhes permitia o controle da matéria-prima e do mercado, além de outras vantagens. A característica principal desse tipo de organização de empresas é sua capacidade de influência e poder sobre uma ampla e extensa abrangência geográfica, cultural e social, além das possibilidades de obtenção de maiores lucros. Já naquela época, as corporações foram acusadas de fixação de preços, exploração e outros abusos, resultando em uma reação política do governo norte-americano, que instituiu novas formas de regulação para esse tipo de organização, no final do século XIX (Clinard, Yager, Brissette, Petrashek, & Harries, 1979).

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

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

 CC BY

Os efeitos das mudanças advindas com a consolidação das corporações dividem as opiniões de estudiosos. Aqueles que defendem os efeitos positivos apontam para o desenvolvimento de novos bens e serviços, preços mais baixos, criação de empregos, melhora no nível de educação formal das pessoas, além da prosperidade das classes de menor poder aquisitivo. Já aqueles que apontam os efeitos negativos criticam a influência que as corporações exercem sobre a vida das pessoas, principalmente pelo seu poder ideológico, que legitima o seu *status quo* como o único e desejável. Grandes negócios implicam o alto poder econômico nas mãos de poucos; empregados têm poder limitado de barganha com seus empregadores e são mais vulneráveis às crises econômicas, além de serem forçados a aceitar condições de trabalho precárias, com longas horas de trabalho e baixos salários (Pearce & Tombs, 1999).

Considerando as influências sociais e econômicas, não se pode esquecer que o Estado, em parte do mundo, deixou de lado seu papel de interventor: empresas públicas foram privatizadas, novos instrumentos financeiros e fiscais facilitaram formas mais eficientes de produção, e as políticas de livre comércio em nível mundial fizeram emergir um sistema globalizado que minou o poder de negociação de regiões e nações (Bauman, 1999). Não é exagerado dizer que o Estado tornou-se um aparato utilizado por um bloco de poder, por exemplo, corporações de uma determinada indústria, para assegurar, manter e ampliar seu domínio sobre a economia, política e sociedade, bem como sobre as agências reguladoras (Pearce & Tombs, 1999). As corporações passaram, então, a gozar do poder da mobilidade, o que lhes permitia escolher os recursos, a mão de obra e a localização de suas operações de modo a obterem condições de produção vantajosas, ao mesmo tempo que eliminariam quaisquer tipos de limitações e restrições.

Quanto às relações entre governos e corporações, essas se valem de conexões políticas, como a doação para campanhas de políticos e a composição do conselho de administração com a presença de membros com *background* em política (Camilo, Marcon, & Bandeira-de-Mello, 2012). Analisando o contexto norte-americano, Barley (2007) mostra o poder de influência das corporações sobre as instituições sociais, inclusive em democracias, contrariando premissas tradicionais de teóricos organizacionais de que apenas o ambiente afeta organizações e essas afetam umas às outras. Barley (2007) mapeia as conexões de lobistas, nos Estados Unidos, representando corporações com políticos norte-americanos para intervirem no ambiente de modo a manter interesses corporativos, influenciando as instituições sociais de três modos: (1) criando uma legislação que favorece cidadãos corporativos; (2) limitando a criação de agências reguladoras que protegem o bem público dos atos das corporações e das externalidades por elas criadas; e (3) privatizando funções que deveriam ser desempenhadas pelos governos local, estadual e federal.

E, assim, o Estado perdeu forças no papel de mediação entre o mercado e a sociedade, o que fez surgir uma "nova proliferação de Estados soberanos frágeis e impotentes" (Bauman, 1999, p. 75), enquanto as multinacionais consolidavam sua influência e poder para obter concessões e instalarem-se em um ambiente de negócios favorecido pela abundância de mão de obra qualificada a um custo baixo, bem como pela baixa regulamentação das condições de trabalho. Essa influência e poder crescem à medida que ocorrem mudanças nas formas de organização das corporações, como uma onda geral de fusões que se iniciou depois da segunda metade do século XX, principalmente no Reino Unido e nos Estados Unidos. Fusões, aquisições, alianças estratégicas e *joint ventures* tornaram-se estratégias comuns para as corporações, pois permitem a elas compartilhar custos e riscos, à medida que aumentam seus lucros, mercado e poder, além de tornarem os políticos impotentes para exercer qualquer controle sobre elas.

Na luta por sobrevivência em um ambiente de competição acirrada por recursos, as corporações, de maneira intencional ou não, de modo indireto ou direto, adotam condutas que podem acarretar crimes que, na maioria das vezes, são levados para os bastidores na vida social. Tais condutas, decisões e ações, no âmbito das grandes corporações, podem constituir-se em crimes contra a sociedade, consumidores, empregados, comunidade, investidores, governos e meio ambiente. Para referir-se a essas ações empresariais, textos jornalísticos e endereços eletrônicos especializados têm propalado a expressão crimes corporativos, um termo que começou a tomar corpo desde a última década do século XX (Erp, 2018).

Russel Mokhiber, jornalista e editor do semanário norte-americano *Corporate Crime Reporter*, fundado em 1987, compilou, em seu livro *Corporate crime and violence: Big business power and the abuse of the public trust*, em 1988, casos expressivos de crimes cometidos por empresas. Mokhiber (1995) traça o perfil de 36 casos de violência e crimes corporativos ocorridos até os anos 1980, com detalhes da violência cometida contra mulheres, crianças, meio ambiente e consumidores, enfatizando as consequências danosas de condutas empresariais. Crimes corporativos têm se tornado cada vez mais comuns, sendo noticiados em jornais sob variadas denominações, deixando claro que são um problema que transcende o nível do indivíduo, estando sua raiz nas estruturas, processos e decisões corporativas (Erp, 2018).

Os crimes corporativos são largamente discutidos no âmbito da Sociologia, do Direito e da Economia, todavia as abordagens são variadas, e nem sempre as suas origens estão associadas à conduta criminosa. Em relação à conceituação do termo crime corporativo, há uma profusão de nomenclaturas, entre elas a expressão *white collar crime* (e. g. Sutherland, 1940), considerada uma das primeiras referências a crimes cometidos nas *suites*, além dos termos crime ocupacional e crime organizacional.

A literatura especializada sobre crime corporativo o associa ao desempenho anterior da empresa, às pressões e barreiras para se obter desempenho superior, à estrutura, ambiente, tecnologia e outras variáveis organizacionais. Além disso, o princípio da autonomia assegura que os proprietários e acionistas nunca arcarão com os custos dos prejuízos sociais causados por corporações (Whyte, 2018). Tampouco as corporações arcam com os custos para evitarem que crimes corporativos aconteçam, preferindo outras práticas que desloquem esses custos para a sociedade em geral, que pouco associa a palavra crime aos acontecimentos do ambiente corporativo, ainda que estes ocorram na busca de alcançar os objetivos corporativos de lucro e desempenho. Contudo, "Como evitar os crimes corporativos?" não é uma questão simples de ser respondida, exigindo esforços para a compreensão, primeiro, da dimensão em que eles ocorrem na sociedade e, segundo, das inter-relações entre os diversos atores envolvidos na sua produção.

A SERIEDADE DOS CRIMES CORPORATIVOS

Na literatura sobre crimes corporativos, um aspecto que chama a atenção é a opinião pública sobre o assunto: os crimes comuns (ou de rua) assustam mais do que os crimes corporativos, mesmo que esses acarretem um conjunto de danos e prejuízos que vão além daqueles, atingindo proporções inestimáveis, haja vista que as estatísticas oficiais quanto a esses custos não levam em conta doenças provocadas pela poluição ambiental, a venda de produtos que provocam danos aos consumidores, as condições de trabalho potencialmente perigosas por exposição a produtos tóxicos, entre outros, sendo esses custos, portanto, subdimensionados. Além das dificuldades contextuais para identificar e mensurar os custos e as consequências desses crimes, em muitos destes não são relatados os prejuízos reais para não constranger os negócios envolvidos.

De modo geral, sem considerar contextos específicos de violência, os custos financeiros do *white collar crime* são tão altos ou maiores do que os custos daqueles considerados "crimes comuns" ou crimes de rua (Cohen, 2015). Porém, os prejuízos das relações sociais vão além das perdas financeiras provocadas por aquele tipo de crime. Isso porque o *white collar crime* viola a confiança, gerando baixa moral social, além de produzir desorganização social em larga escala, o que não ocorre com os crimes nas ruas, cujos efeitos sobre as instituições e organizações sociais são menores.

De fato, comparativamente, os altos custos dos crimes corporativos excedem, e muito, os custos do crime individual, visto que, nos primeiros, um simples e único ato pode resultar em milhares de vítimas. A violação de normas de segurança no trabalho pode acarretar várias mortes e acidentes; a contaminação e poluição ambiental podem afetar diversas famílias e comunidades; a utilização de materiais perigosos na produção de bens pode aumentar o risco de problemas de saúde de muitos trabalhadores e consumidores. Além disso, as vítimas não se reduzem aos indivíduos apenas, mas, também, aos pequenos negócios que podem ser levados à insolvência.

A discussão sobre a responsabilização e culpabilidade das corporações pelos danos causados pelas suas atividades recai em três pontos principais: o primeiro diz respeito ao fato de que uma corporação jamais poderá ser presa; o segundo refere-se ao fato de que reconhecer que, se as corporações estão sujeitas às leis criminais tal qual o indivíduo, isso pode significar que estas têm os mesmos direitos e responsabilidades; e o terceiro reside nas atitudes de tolerância ou nas reações da sociedade em relação à conduta das organizações. Esse terceiro aspecto decorre, principalmente, pela exaltação do mercado e da iniciativa privada como responsáveis pelo progresso e desenvolvimento econômico das nações, o que levou a uma santificação das corporações.

Sutherland (1940) argumentava que o baixo número de condenações dos crimes corporativos na justiça criminal dos Estados Unidos devia-se, em parte, à ausência de sanções penais efetivas para esse tipo de crime. Essa ausência é explicada pelo autor pela impossibilidade de condenar a corporação à morte ou à prisão, sendo a única penalidade possível o pagamento de multa, a qual, na verdade, é arcada pelos acionistas na forma de redução de seus dividendos. Outro argumento do autor é que o *white collar crime* encontrou espaço para crescer em virtude da tolerância do público em relação a sua ocorrência, o que tem se modificado ao longo desses anos.

É fato que a opinião pública desempenha papel fundamental no debate sobre a culpabilidade da corporação, visto sua influência no controle das ilegalidades no mundo dos negócios. A população em geral considera as ofensas corporativas como sérias apenas quando suas consequências são físicas,

 CC BY © RAE | São Paulo | **V. 59** | **n. 6** | **nov-dez 2019** | 435-441 ISSN 0034-7590; eISSN 2178-938X

substanciais e, relativamente, imediatas. Todavia, o *white collar crime* é um crime real e, mesmo que não seja denominado de crime, ordinariamente, isso não o torna menor. A opinião pública desempenha, ainda, papel relevante na regulação e controle dos crimes corporativos, tanto que a ambiguidade da opinião pública diante do comportamento corporativo ilegal faz com que a lei também seja ambígua.

Essa questão foi explorada em uma pesquisa (Unnever, Benson, & Cullen, 2008) realizada com norte-americanos para identificar se estes desejavam aprovar regulamentos mais rigorosos do mercado de ações e defender sanções penais mais punitivas para os executivos que ocultam a verdadeira condição financeira da empresa. Conforme esses autores reforçam, os sentimentos do público são potencialmente importantes na formação de políticas de controle dos crimes, especialmente se esses ocorrem quando a mensagem do público é que algo deve ser feito para coibir os crimes corporativos.

No contexto norte-americano, a criminalidade nas ruas, desde os anos 1980, recebeu maior atenção por parte do governo, resultando na adoção de políticas públicas de controle mais punitivas contra o crime. Porém, em relação à criminalidade corporativa, a despeito da onda de escândalos corporativos no país, não foi dirigida a atenção necessária, o que resultou em uma lacuna a ser preenchida. As violações corporativas, em sua maioria, eram tratadas por meio de procedimentos civis e regulatórios, e não na esfera da justiça criminal, o que confere a essas violações o caráter de um evento acidental.

Algumas razões justificam esse tratamento dispensado às violações das leis por parte das corporações. A primeira delas é a crença, entre os economistas e os formuladores de políticas, de que o livre mercado se autorregula, não havendo necessidade de lei criminal para controlar o comportamento prejudicial das corporações. Em um ponto de vista oposto, estão autores como Snider (1990), que acreditam que os controles do governo, na forma de regulações ou leis criminais, são necessários e apropriados.

O movimento contra os crimes corporativos, nos EUA, ocorreu a partir de 1970, resultando em um incremento no uso da lei criminal contra as corporações. Unnever et al. (2008) analisam a evolução da regulação e legislação criminal dirigida ao mundo corporativo, nos Estados Unidos, em três passos: (1) um tipo de escândalo é descoberto e (2) o clamor público demanda uma ação do governo, o qual reage, (3) formalizando uma acusação ou criando novas leis e regulações.

A reação da população em relação às condutas das empresas é relevante para que se estabeleçam políticas públicas para a criminalização do crime corporativo, pois a legitimidade do sistema de justiça criminal de um país pode ser colocada à prova caso o governo falhe em responder a crimes corporativos de grande repercussão e que resultam em danos para um contingente significativo de vítimas.

As instituições e corporações presentes na sociedade refletem as desigualdades econômicas e sociais, produzindo práticas e padrões transgressores, os quais, por sua vez, resultam em processos normativos econômicos e políticos dessa sociedade. Nesse sentido, uma visão crítica do crime corporativo pode reconhecer que o crime, assim como um processo social, é moldado por aqueles que desfrutam de poder e influência político-econômica de modo a garantir que a denominação do que seja crime vá refletir sua visão de mundo e seus interesses, no caso das corporações, poder econômico, social e político.

PODEMOS FALAR EM GENOCÍDIOS?

As ações criminosas das corporações ganharam status de genocídios (Brook, 1998; Kelly, 2013, 2016; Medeiros & Alcadipani, 2018; Stokes & Gabriel, 2010), e as organizações e sua racionalidade instrumental também foram associadas ao Holocausto (Bauman, 1998; Black, 2001). A despeito das diferentes significações que o termo adquiriu ao longo da história, bem como as controvérsias que pairam sobre ele, não se trata de exagero fazer tais associações, pois genocídio é definido pelo assassinato em massa de maneira organizada e, mesmo sendo feito pelo Estado e milícias, as corporações são cúmplices (Stel, 2014), na forma do provimento de armas e outros recursos. Embora extermínios em massa tenham ocorrido antes, foi somente no século XX que o termo genocídio foi cunhado pelo jurista polonês e judeu Raphael Lemkin, em 1943, no esforço para denunciar as atrocidades nazistas, então refugiado nos Estados Unidos da América. Sem dúvida, Lemkin foi influenciado pelo Holocausto, que vitimou muitos membros de sua família, para definir, em seu livro *Axis rule in occupied Europe* (1944), o genocídio como assassinato de grupos étnicos, religiosos e nacionais (Naimark, 2015).

Para além dos crimes corporativos que levam diretamente centenas de pessoas à morte, empresas do setor financeiro agem em cumplicidade com governos e outras organizações que praticam genocídios, como o BNP Paribas, que é acusado de financiar a compra de armas utilizadas para cometer genocídio em Ruanda, em 1994 (BBC NEW, 2014). Outro exemplo é a colaboração da IBM e da Ford com o Holocausto (Lima, 2016). A cumplicidade com o genocídio, que pode se configurar de diferentes formas, é um ato passível de punição no artigo 3º. da Convenção para a Prevenção e Punição do Crime de Genocídio, documento internacional elaborado depois da II Guerra Mundial com o objetivo

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

de proteger a pessoa humana dos genocídios nazistas. Nesse mesmo documento, genocídio é definido pelos atos cometidos para destruir um grupo nacional, étnico, racial ou religioso, seja em partes ou no todo. Esse conceito é alvo de polêmicas por não incluir grupos políticos, econômicos e culturais, os quais são excluídos da definição de modo proposital (Schabas, 2009). Também são excluídas as vidas dispensadas pelo capitalismo contemporâneo, como as mortes provocadas pela busca desenfreada pelo lucro econômico, o que Banerjee (2008) denomina necrocapitalismo.

Dessa maneira, a destruição causada pelos crimes corporativos também ocorre em larga escala, assim como nos genocídios. A campanha *No Business With Genocide* foi criada, em 2017, com o propósito de que as corporações não fizessem negócios com regimes que se envolvessem com genocídios ou crimes contra a humanidade. Mokhiber (1995), ao argumentar que o crime corporativo é um crime mais violento do que os crimes de rua, sustenta seus argumentos por estatísticas mundiais e norte-americanas sobre pessoas assassinadas nas ruas nos Estados Unidos e no mundo e aquelas que morrem, no mesmo período, em virtude de doenças ocupacionais e falta de segurança no trabalho, além das vítimas de produtos prejudiciais à saúde.

A Union Carbide provocou mais de 3 mil mortes, além de deixar mais de 50 mil pessoas incapacitadas para trabalhar, no caso de Bhopal, na Índia, em 1984. O setor de mineração causa impactos nos direitos humanos e no meio ambiente, destruindo vidas. *Vide* a condenação de uma mineradora a pagar indenização pelos danos ambientais e sociais em virtude de ter provocado a contaminação por chumbo, atingindo mais de 18 mil pessoas, na cidade de Santo Amaro da Purificação, na Bahia. Sete milhões de fumantes morrem por ano. O trabalho escravo contemporâneo atinge 46 milhões de pessoas. A indústria bélica segue fornecendo armas para conflitos, produzindo bilhões de balas e milhões de armas cada vez mais sofisticadas, com um gasto de um trilhão e meio de dólares em armamento, enquanto uma pessoa é atingida por uma bala por segundo. E centenas de trabalhadores do setor de vestuário, em Bangladesh, Dhaka e Chittagong, morreram nos desabamentos dos edifícios em que trabalhavam para fornecedores de grandes marcas.

Não se pode precisar os números de vidas ceifadas e dos prejuízos materiais e sociais, o que é difícil identificar e mensurar em crime corporativos. Além de todas as mortes físicas de pessoas, rios, fauna e flora, há de se considerar a morte psicológica daqueles que sobrevivem e a insegurança da população acerca de informações que vieram à tona sobre a fragilidade dos sistemas operacionais das empresas que cometem crimes corporativos, e que antes eram guardadas em segredo.

POR QUE AS ORGANIZAÇÕES SE TORNAM CRIMINOSAS

O comportamento criminal no âmbito das corporações não deve ser analisado como um desvio pessoal, mas, sim, como um produto das relações e interações humanas em contextos específicos, a depender de suas características econômicas, legais, organizacionais e normativas, pois, independentemente do grau de motivação pessoal dos envolvidos, a criminalidade corporativa constitui-se em uma atividade legítima do contexto em que emerge.

Não é simples dizer o porquê de as organizações tornarem-se criminosas. Talvez essa resposta vá ser encontrada a partir de estudos com abordagens inter e multidisciplinares, em suas diferentes perspectivas e disciplinas. Para além dos resultados de pesquisas já realizadas sobre as causas da criminalidade corporativa, os quais apontam fatores organizacionais, institucionais, a cultura organizacional, pressões por resultados e restrições econômicas, entre outros, nós endereçamos aqui reflexões sobre como os crimes corporativos são organizados pelas corporações.

A primeira delas diz respeito à participação de agências do governo de duas formas: quando corporações contratadas pelo governo engajam em práticas desviantes ou têm a aprovação do governo para tal; e quando as instituições regulatórias governamentais falham em restringir as atividades de negócios desviantes. Trata-se do chamado *state-corporate crime*, um tipo que compreende a intersecção entre governos e corporações para produzir danos sociais, tendo o termo sido apresentado, pela primeira vez, por Richard Kramer, em 1990, durante o encontro anual da Society for the Study of Social Problems.

Outra reflexão é sobre a disposição que as corporações têm, naturalmente, para cometer crimes ou transgredir a lei no interesse da maximização dos lucros (Tombs & Whyte, 2015). As organizações provêm oportunidades para a mobilização de conhecimento necessário para cometer crimes, por exemplo: guardando segredos, disfarçando ilegalidades, omitindo lucros ilegais, destruindo evidências, pagando para legisladores, políticos e autoridades para que a lei não seja aplicada a elas. É possível concluir, então, que as organizações são uma arma para cometer crimes corporativos, e, ainda, organizações colaboram entre si para cometer crimes corporativos, seja por meio de *joint ventures* ou outras alianças estratégicas.

Ainda, há de se considerar a conduta dos dirigentes das organizações, responsáveis por decisões que levam aos crimes corporativos, na busca de atingir os objetivos corporativos. Desse modo, as motivações relacionadas a fatores econômicos, às decisões dos gestores/executivos e às relações estabelecidas com

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

o Estado atuam de maneira interligada para que as organizações cometam crimes e, posteriormente, sejam reincidentes.

As corporações, então, não se tornam criminosas. Se elas são o motor do capitalismo contemporâneo, e o lucro é o principal objetivo, algum crime, potencialmente, ocorrerá. Indenizações, quando pagas, não impactam os lucros. Veja-se que, na maioria dos casos, mesmo que as ações da empresa que cometeu o crime corporativo tenham tido uma baixa, e apesar de ter prejuízos, depois de certo tempo, a empresa consegue retomar seus lucros e pagar generosos bônus aos seus executivos.

BASES PARA REFLEXÕES SOBRE O CONTROLE DOS CRIMES CORPORATIVOS

Esse cenário traçado é assustador: o domínio das corporações sobre nossas vidas torna impossível a tarefa de desmantelar a potencialidade de crimes corporativos, cuja letalidade pode ser comparada à de um genocídio. Não temos uma resposta para a simples questão de "o que fazer?". Talvez algo que se deva ter em mente é que crimes corporativos não devem ganhar *status* de negligência corporativa e ser julgados como tal, ou seja, pela lei civil. Em casos como esses, as decisões judiciais acabam em multas pagas ao governo, ou, em alguns casos, em restituições às partes lesadas. Rejeitamos a ideia de que o Estado não deva intervir por meio de regulamentações, proibições e punições severas para as corporações e seus dirigentes. E, sim, concordamos com intervenções que impliquem a perda de autonomia desse modelo de negócio, o que significa também a retirada dos privilégios que isentam os proprietários de arcar com todos os danos que causam (Whyte, 2018).

Mudanças nos tipos de lideranças políticas são necessárias, principalmente no que diz respeito ao financiamento ou doações de campanhas políticas por parte das corporações e empresas. Para controlar os abusos das corporações, é necessária a emancipação dos políticos quanto a esses vínculos, cujo preço a se pagar é alto, muitas vezes, com vidas.

As conexões das corporações com os legisladores são empecilhos para que não tenhamos mais que lamentar pelas vítimas de crimes corporativos. Há quase 30 anos, Sargent (1990) já alertava para a necessidade de problematizar a criminalização das violações das corporações e chamava atenção, ainda, para que esforços fossem dirigidos para explorar outras vias para o controle do crime corporativo. Assim, espera-se que a sociedade, na pessoa de seus cidadãos, questione as práticas e condutas das corporações que afetam o modo de vida atual e futuro. As mudanças sociais requeridas para que os crimes corporativos sejam evitados, ou melhor, para que não tenhamos que arcar com os danos e prejuízos provocados por eles, que são lucrativos para as corporações, têm sua gênese no questionamento quanto à dominação das corporações na sociedade contemporânea, bem como sobre suas responsabilidades diante da transgressão corporativa.

A corporação como modelo de negócios constitui-se na maior força do capitalismo contemporâneo. O princípio da separação corporativa, o qual prevê que as responsabilidades da corporação e a propriedade dos seus ativos são exclusivos da corporação, portanto, os acionistas não têm responsabilidades sobre seus crimes e má conduta, é um incentivo para que as decisões corporativas não sejam tomadas com base nas suas consequências danosas. Nenhuma sociedade quer lamentar por vidas perdidas, nem quer mais reivindicar indenizações por suas perdas materiais ou derramar lágrimas pela destruição de suas identidades construídas, nem mesmo reviver lembranças traumáticas. O que a sociedade deseja é que crimes corporativos não mais aconteçam. E, no nosso entendimento, a dissolução dessa forma de organizar os negócios, a corporação, bem como de outras que facilitem a extrema acumulação de riqueza e poder, seria uma via para que se evitem os crimes corporativos.

Encerramos esta pensata com alguns pontos que devem ser considerados como base para nossas reflexões sobre o controle de crimes corporativos, pois entendemos que estes provocam danos e prejuízos com dimensões imensuráveis e custos inestimáveis, incluindo-se vidas, o que não tem preço. Não se trata de propor soluções para um problema social complexo como a atuação criminosa de corporações, mas, sim, apresentar bases para uma reflexão analítica sobre os crimes corporativos como genocídios.

NOTA DE AGRADECIMENTO

Este artigo faz parte dos resultados da pesquisa realizada durante o estágio pós-doutoral em Administração na Fundação Getulio Vargas, Escola de Administração de Empresas de São Paulo. Esta pesquisa contou com o apoio do CNPq.

Os autores agradecem aos revisores pela leitura cuidadosa e sugestões que contribuíram sobremaneira para o aprimoramento da versão inicial.

REFERÊNCIAS

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 ISSN 0034-7590; eISSN 2178-938X