

Big Data em Marketing: Análise Bibliométrica e Discussão dos Principais Estudos Desenvolvidos na Área

Autoria

Claudio Sunao Saito - cssaito@uol.com.br

Prog de Mestr e Dout em Admin em Gestão Internacional/Dout e Mestr em Admin - PMDGI/ESPM - Esc Sup de Prop e
MKT de São Paulo/Ass Esc Sup de Prop e MKT

Marielza Rita Cavallari - marielza.cavallari@espm.br

Graduação/ESPM - Esc Sup de Prop e MKT de São Paulo/Ass Esc Sup de Prop e MKT

Vivian Iara Strehlau - vstrehlau@gmail.com

Prog de Mestr Prof em Comportamento do Consumidor - MPCC/ESPM - Esc Sup de Prop e MKT de São Paulo/Ass Esc
Sup de Prop e MKT

Resumo

Este artigo tem como objetivo principal a investigação da literatura que aborda o Big Data na área de marketing. Com o uso de técnicas, ferramentas e indicadores bibliométricos são identificados, classificados, agrupados os principais artigos que abordam o tema, possibilitando o estudo da forma como o Big Data tem contribuído ou impactado teorias e práticas na área de marketing. Os resultados da pesquisa indicam que a maioria dos estudos de Big Data em marketing focam em métodos, técnicas e uso de ferramentas de mineração e análise de dados com base em registros de transações comerciais e avaliações realizadas por consumidores e clientes. Foram encontrados, em menor número, artigos com estudos baseados em informações obtidas a partir do Big Data e também revisões anteriores sobre o assunto na área de marketing.



Big Data em Marketing: Análise Bibliométrica e Discussão dos Principais Estudos Desenvolvidos na Área

RESUMO

Este artigo tem como objetivo principal a investigação da literatura que aborda o *Big Data* na área de marketing. Com o uso de técnicas, ferramentas e indicadores bibliométricos são identificados, classificados, agrupados os principais artigos que abordam o tema, possibilitando o estudo da forma como o *Big Data* tem contribuído ou impactado teorias e práticas na área de marketing. Os resultados da pesquisa indicam que a maioria dos estudos de *Big Data* em marketing focam em métodos, técnicas e uso de ferramentas de mineração e análise de dados com base em registros de transações comerciais e avaliações realizadas por consumidores e clientes. Foram encontrados, em menor número, artigos com estudos baseados em informações obtidas a partir do *Big Data* e também revisões anteriores sobre o assunto na área de marketing.

Palavras-chave: Big Data; Marketing; Mining; Analytics; Scopus

1 Introdução

O *Big Data* é um assunto muito discutido na atualidade, principalmente devido às possibilidades de acesso a dados oferecidas pelas tecnologias de e-commerce e redes sociais. Embora exista alguma relatividade na avaliação da quantidade de dados que pode ser considerada “grande”, um artigo de Cox e Ellsworth (1997) demonstra que discussões sobre o assunto já ocorriam na National Aeronautics and Space Administration (NASA) em 1997, quando afirmam que o problema do *Big Data* é a sobrecarga da memória, do disco local e do disco remoto dos sistemas de computação.

A relação entre o termo *Big Data* e a área de Marketing é recente. Nas bases de dados Web of Science e Scopus, o primeiro documento onde constam as palavras-chave “Big Data” e “marketing” em conjunto, foi o artigo apresentado por Irudeen e Samaraweera (2013) no International Conference on Advances in ICT for Emerging Regions (ICTer) e o primeiro documento onde constam os termos nas palavras-chave foi o artigo publicado por Banerjee et al. (2013).

Este estudo tem como objetivo a investigar e revisitar a literatura que aborda o *Big Data* na área de marketing. A partir do uso de técnicas bibliométricas, são identificados, classificados, agrupados e descritos os principais artigos que abordam o tema, possibilitando o entendimento da forma como o *Big Data* tem contribuído ou impactado teorias e práticas na área de Marketing. Nos próximos tópicos serão discutidas as definições de *Big Data*, apresentada a metodologia da pesquisa e os resultados da pesquisa.

2 Big Data

Big Data é um termo que descreve o grande volume de dados (SAS, 2018) ou interpretação de grandes volumes de dados de grande variedade (CANALTECH, 2018). O ponto em comum entre essas definições é a disponibilidade de alto volume de dados. Não por acaso, segundo Francisco (2014), o “big” do *Big Data* se traduz no volume de dados cuja escala, distribuição, diversidade e/ou velocidade de criação requerem o uso de novas tecnologias de armazenamento e análise para permitir a captura do valor inserido neles.

Haider e Gandomi (2015) destacam ainda que as definições de volume em *Big Data* são relativas e variam de acordo com fatores, como o tempo e o tipo de dados. O que pode ser considerado *Big Data* hoje, pode não atingir o limite no futuro, porque as capacidades de

armazenamento aumentarão, permitindo que conjuntos de dados ainda maiores sejam capturados. Além disso, o tipo de dado, define o que significa "grande". Dois conjuntos de dados do mesmo tamanho podem exigir diferentes tecnologias de gerenciamento de dados de acordo com o tipo, por exemplo, dados tabulares versus dados de vídeo.

Então, o “big” do *Big Data* envolve mais do que o volume, devendo abranger outras perspectivas além da quantidade de dados. Para Laney (2001), além do volume, variedade e velocidade são as três dimensões (ou três “Vs”) que estabelecem o desafio da gestão de dados. A perspectiva da velocidade, envolve a capacidade dos dados serem cada vez mais assíncronos e real time. Com os recursos tecnológicos atuais, dados referentes às organizações e pessoas são registrados no momento em que ocorrem, podendo ser acessadas no momento em que os fatos são gerados. A terceira perspectiva, é relativa à variedade de dados, sendo elas estruturadas ou não estruturadas. Enquanto os dados estruturados são coletados e armazenados, em geral, pelos ERPs, os não estruturados podem ser acessados em diversas fontes e formatos. Segundo Francisco (2014), de 80% a 90% dos dados da internet não são estruturados — páginas “www” em formato HTML ou XML, dados de *clickstream*, fotos, imagens, vídeos, textos em linguagem natural, mapas etc. Para o autor, dados de redes sociais, que compreendem atualmente um dos tipos de dados não estruturados mais analisados pelas empresas, podem ser utilizados, por exemplo, para avaliação de uma determinada organização, marca ou um evento específico.

A disponibilidade de um grande volume de dados, atualizados, variados e confiáveis, não garante, porém, a criação de valor para as organizações. Para isso, torna-se necessária a mobilização de recursos que permitam o acesso, organização, classificação e análise das informações — essa última conhecida como *Analytics* ou *Big Data Analytics*. Para ajudar a organizar as informações, modelos estatísticos são aprimorados para incluir informações de localização, sendo essa uma tendência do *Big Data Analytics*: o uso de técnicas de estatística espacial. Segundo Francisco (2014), muitos negócios podem se beneficiar dessas técnicas: modelos tradicionais de previsão de renda e de crédito, no segmento bancário e no mercado financeiro em geral; modelos de gestão de risco e desempenho, que qualificam e analisam influências locais no aumento de risco para seguradoras; e o mercado imobiliário, para a correta gestão de compra e venda de ativos e o mapeamento de regiões, visando a melhor precificação de imóveis, em um contexto em que a dinâmica urbana territorial é essencial.

3 Metodologia da pesquisa

Nesse estudo foram utilizadas técnicas, ferramentas e indicadores bibliométricos. Cronin (2001) afirma que estudos bibliométricos tradicionalmente concentraram esforços no acompanhamento de indicadores altamente visíveis e objetivos de atividade acadêmica: publicações e citações. Os indicadores bibliométricos mais utilizados são os indicadores do desempenho científico de organizações, agências e países com base nas contagens de publicações e citações na literatura científica (Narin & Hamilton, 1996). Em relação ao método, Wallin (2005) afirma que estudos bibliométricos incluem padrões de publicação de estudos, *bibliographing* (contagem do número de resumos e indexadores bibliográficos ou bancos de dados que registam o conteúdo da revista em questão), acoplamento bibliográfico (co-citação e co-ocorrência) e análise de citações de artigos e patentes científicas. Os métodos bibliométricos são quantitativos por natureza, mas são usados para fazer pronunciamentos sobre características qualitativas. O principal objetivo de todos os tipos de exercícios bibliométricos é transformar algo intangível em uma entidade gerenciável (Wallin, 2005).

Na pesquisa de artigos relacionados ao tema, foi utilizada a base de dados bibliográficos Elsevier Scopus. A base Scopus se assemelha ao tradicional Thomson ISI Web of Science no desenvolvimento da coleção e na cobertura mundial (Moya-Anegón et al., 2007), mas oferece

No software estatístico SciMat (Cobo *et al.*, 2012), foram gerados 255 agrupamentos baseados em co-ocorrência de palavras-chave. Com base na análise de conteúdo, três clusters desses clusters foram selecionados para a pesquisa bibliométrica: “consumer”, “costumer” e “marketing”. Esses clusters foram selecionados por conter palavras-chave de artigos diretamente relacionados ao tema da pesquisa. Os artigos provenientes desses clusters foram identificados e, após eliminação de duplicidade, resultaram em uma amostra de 37 artigos a serem analisados.

4 Resultados da pesquisa

Nos tópicos a seguir, serão apresentados os resultados da análise bibliométrica e da análise de conteúdo dos artigos provenientes da amostra escolhida.

4.1 Análise bibliométrica

Com base na amostra de 37 artigos, foram realizadas análises e levantamento de indicadores bibliométricos para maior entendimento do tema abordado no estudo. Em relação ao ano das publicações, verifica-se que esse assunto é recente, sendo x% dos artigos publicados a partir de 2014. Os dados relativos aos anos de publicação dos artigos estão apresentados na tabela 1 e figura 2.

Tabela 1

Número de publicações em *journals* por ano

Ano	n	%
2013	1	3
2014	7	19
2015	6	16
2016	10	27
2017	5	14
2018	8	22
Total	37	100

Fonte: Elaborado pelos autores com utilização de dados da base Scopus

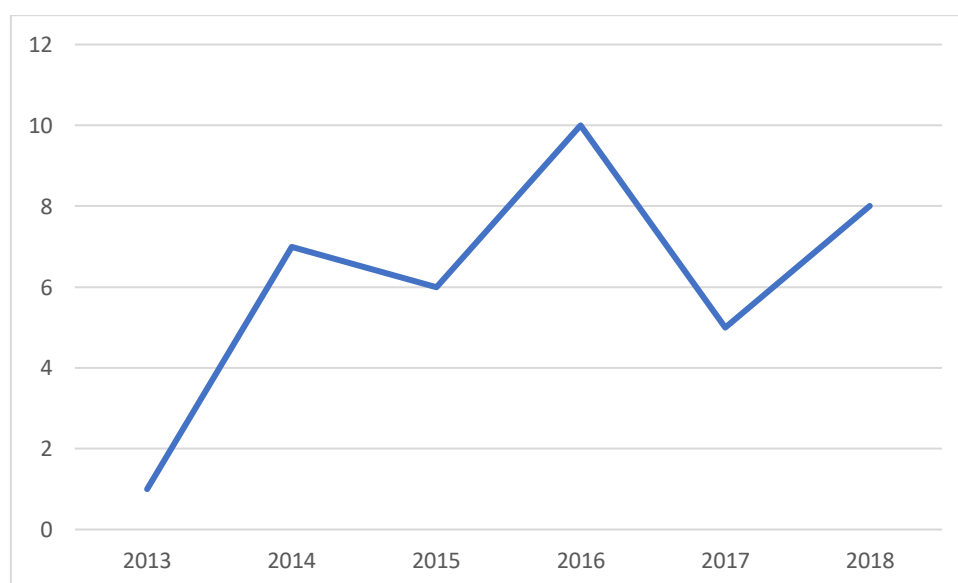


Fig. 2 Crescimento do número de publicações ao longo do tempo.

Fonte: Elaborado pelo autor com dados da base Scopus.

Em relação à amostra coletada, os *journals* com maior número de publicações foram o Journal of Marketing (H-index 208), com 3 artigos e o Marketing Science (H-index 108), com 3 artigos publicados, ambos classificados pelo SCImago Journal & Country Rank – SJR(2018) no Q1 (melhor quartil) nas áreas de Business and International Management e Marketing. Na tabela 2, estão apresentados os *journals* com uma ou duas publicações de artigos da amostra coletada. Verifica-se que a publicação de artigos sobre marketing e Big Data é pulverizada, não existindo concentração de publicações em journals específicos.

Tabela 2

Número de artigos publicados por Journal

Fonte	n	%
Journal of Marketing	3	8
Marketing Science	3	8
Electronic Commerce Research and Applications	2	5
Industrial Management and Data Systems	2	5
International Journal of Production Research	2	5
Journal of Direct Data and Digital Marketing Practice	2	5
Technological Forecasting and Social Change	2	5
Annals Of Tourism Research	1	3
Business Horizons	1	3
Espacios	1	3
European Management Journal	1	3
European Research on Management and Business Economics	1	3
Information And Management	1	3
International Journal of Electronic Customer Relationship Management	1	3
International Small Business Journal Researching Entrepreneurship	1	3
Journal of Arts Management Law And Society	1	3
Journal of Business Research	1	3
Journal of Cleaner Production	1	3
Journal of Consumer Marketing	1	3
Journal of Interactive Marketing	1	3
Journal of Marketing Analytics	1	3
Journal of Marketing Research	1	3
Journal of Service Research	1	3
Journal of Services Marketing	1	3
Journal of Theoretical and Applied Electronic Commerce Research	1	3
Quality Access To Success	1	3
Tourism Management	1	3
Wseas Transactions on Business And Economics	1	3
Total	37	100

Fonte: Elaborado pelos autores com utilização de dados da base Scopus

De forma similar, não existe a concentração de publicações por instituições de ensino. Destacam-se na amostra as universidades de Houston e Maryland com 3 publicações cada uma e as universidades Northwestern, Chinese Academy of Sciences, San Francisco State e a Northeastern com duas publicações. Verifica-se na tabela 4, que 84% dos artigos publicados apresentam filiação a universidades distintas.

Tab. 4

Artigos publicados por instituição de ensino.

Instituições	n	%
University of Houston	3	4
Northwestern University	2	2
Chinese Academy of Sciences	2	2
University of Maryland	2	2
San Francisco State University	2	2
Northeastern University	2	2
Outras	71	85
Total	84	100

Fonte: Elaborado pelos autores com utilização de dados da base Scopus

Na tabela 5, são apresentados os artigos que compõem a amostra analisada, ordenadas por quantidade de citações (n). Para Bornmann et al. (2008), citações são indicadores de impacto da pesquisa, ou seja, o grau em que ela tem sido útil para outros pesquisadores e Van Raan (2004) afirma que a análise bibliométrica baseada em quantidade de citações fornece indicadores de impacto e influência internacional.

Tabela 5

Artigos da amostra por ordem de citação

Autores	Título do Artigo	Ano	n	%
Erevelles S., Fukawa N., Swayne L.	Big Data consumer analytics and the transformation of marketing	2016	89	13
Leeflang P.S.H. et al.	Challenges and solutions for marketing in a digital era	2014	82	12
Tirunillai S., Tellis G.J.	Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation	2014	80	12
Rust R.T., Huang M.-H.	The service revolution and the transformation of marketing science	2014	71	11
He W., Wu H., Yan G., Akula V., Shen J.	A novel social media competitive analytics framework with sentiment benchmarks	2015	54	8
Wedel M., Kannan P.K.	Marketing analytics for data-rich environments	2016	44	7
Dolnicar S., Ring A.	Tourism marketing research: Past, present and future	2014	37	6
Bone S.A., Fombelle P.W., Ray K.R., Lemon K.N.	How Customer Participation in B2B Peer-to-Peer Problem-Solving Communities Influences the Need for Traditional Customer Service	2015	24	4
Jun S.-P., Park D.-H., Yeom J.	The possibility of using search traffic information to explore consumer product attitudes and forecast consumer preference	2014	22	3
Martin K.D., Borah A.	Data privacy: Effects on customer and firm performance	2017	21	3

Palmatier R.W.				
Culotta A., Cutler J.	Mining brand perceptions from twitter social networks	2016	18	3
Hofacker C.F.et al.	Big Data and consumer behavior: imminent opportunities	2016	17	3
Weinberg et al.	Perspectives on big data	2013	16	2
Jin J et al.	Understanding big consumer opinion data for market-driven product design	2016	15	2
Du R.Y., Hu Y., Damangir S.	Leveraging trends in online searches for product features in market response modeling	2015	14	2
Qin R., Yuan Y., Wang F.-Y.	Exploring the optimal granularity for market segmentation in RTB advertising via computational experiment approach	2017	11	2
Liu Y.-P., Guo J.-F., Fan Y.	A big data study on emitting companies' performance in the first two phases of the European Union Emission Trading Scheme	2017	8	1
Anshari M. et al.	Social customer relationship management and student empowerment in online learning systems	2015	8	1
Amado et al.	Research trends on Big Data in Marketing: A text mining and topic modeling-based literature analysis	2018	7	1
Bendle N.T., Wang X.S.	Uncovering the message from the mess of big data	2016	7	1
Donnelly C., et al.	Digital loyalty card 'big data' and small business marketing: Formal versus informal or complementary?	2015	7	1
Hyrnsalmi S. et al.	Busting myths of electronic word of mouth: The relationship between customer ratings and the sales of mobile applications	2015	5	1
France S.L., Ghose S.	An analysis and visualization methodology for identifying and testing market structure	2016	4	1
Chen Q., Zhang M., Zhao X.	Analysing customer behaviour in mobile app usage	2017	3	0
Rojas Berrio S.P. et al.	Big data: Trend emerging from research in marketing	2016	2	0
Shimomura Y. et al.	A method for identifying customer orientations and requirements for product-service systems design	2018	1	0
George M., Wakefield K.L.	Modeling the consumer journey for membership services	2018	1	0
Talón-Ballester P. et al.	Using big data from Customer Relationship Management information systems to determine the client profile in the hotel sector	2018	0	0
Chiang L.-L.L., Yang C.-S.	Does country-of-origin brand personality generate retail customer lifetime value? A Big Data analytics approach	2018	0	0

Damangir S., Du R.Y., Hu Y.	Uncovering Patterns of Product Co-consideration: A Case Study of Online Vehicle Price Quote Request Data	2018	0	0
Tian J., Zhang Y., Zhang C.	Predicting consumer variety-seeking through weather data analytics	2018	0	0
Ćurko K., Merkaš Z., Silović T.	Challenges of application of the big data in marketing: Case study Croatia	2018	0	0
Verma N., Singh J.	An intelligent approach to Big Data analytics for sustainable retail environment using Apriori-MapReduce framework	2017	0	0
Mylan C.	Gecko improves customer communications for Eurocamp	2016	0	0
Dinu et al.	Studying the consumer behavior through big data	2016	0	0
Stevens L.K.	Making the case: Using a living and historic cultural universe big data system as a knowledge-based foundation for sustaining and recapitalizing the cultural sector	2014	0	0
Shaw R.	The marketing data space race	2014	0	0
Total de citações			668	100

Fonte: Elaborado pelos autores com utilização de dados da base Scopus

4.2 Análise de conteúdo dos artigos da amostra

Os 37 artigos da amostra foram analisados através da técnica de análise de conteúdo (Bardin, 2004), sendo agrupados por assunto e interpretados de maneira a gerar inferências que podem servir como base para discussão e/ou geração de novas teorias. Os principais agrupamentos por assuntos obtidos e os conteúdos dos artigos são apresentados a seguir.

4.2.1 Pesquisas realizadas com uso de *Big Data*

Os artigos desse grupo utilizam o *Big Data* para a execução de pesquisas em assuntos ligados ao marketing.

Hyrnsalmi S. et al. (2015) testaram a relação entre o boca-a-boca eletrônico e a venda de aplicativos em um ecossistema de aplicativos para dispositivos móveis (Google Play) com um conjunto de dados abrangente (mais de 260 milhões de avaliações de clientes; 18 meses), mostrando que valores mais altos de avaliações de clientes estão estatisticamente correlacionados com vendas mais altas.

Utilizando registros de compra de cervejas da base de dados de um grande varejista de Hong Kong, Chiang e Yang (2018) investigaram a influência da relação entre os traços de personalidade do consumidor e os traços do país de origem de marcas de cerveja no valor vitalício do cliente (CLV). Os autores concluíram que consumidores tendem a comprar e adquirir marcas com características semelhantes às suas próprias características de personalidade (ou seja, Japão - tranquilidade, Bélgica - abertura, Irlanda - excitação, etc.).

Damangir, Du e Hu (2018) usaram “um enorme conjunto de dados de solicitações de cotação de preços on-line de compradores de automóveis” para descobrir, probabilisticamente, os padrões predominantes de co-considerações no mercado de automóveis dos EUA.

Com base nos dados obtidos na base Community Independent Transaction Log (CITL) da União Européia, Liu, Guo e Fan (2018) apresentaram um micro estudo sobre os esforços das

empresas emissoras de CO2 em aumentar os lucros e poupar custos durante as negociações de licenças nas duas primeiras fases do Sistema de Comércio de Emissões da União Europeia (EU ETS). Os esforços são medidos por um fator pós-ação do desempenho comercial, que é construído sobre uma série de variáveis comportamentais e monetárias.

4.2.2 Artigos sobre Mining e Analytics

Nesse grupo são apresentados os artigos que apresentam modelos e técnicas para coleta (*mining*) e análise de dados (*analytics*) gerados pelo *Big Data*.

Ao comparar os resultados da pesquisa de atitude baseada em pesquisas convencionais com os resultados das informações de tráfego de busca (Google Insights), Jun, Park e Yeom (2014) revelam que as informações de tráfego de busca indicam o nível de interesse dos consumidores em relação a um produto, os atributos de produto que estão considerando e a importância de cada atributo. Para eles, a utilização desse tipo de informação não apenas supera as limitações impostas pelo alto custo de realização de pesquisas convencionais, em termos de tempo e dinheiro, mas também ajuda a reduzir distorções causadas por erros conscientes ou inconscientes cometidos pelos entrevistados.

Tirunillai e Tellis (2014) propõem uma estrutura unificada, baseada no modelo probabilístico denominado Latent Dirichlet Allocation (LDA), que utiliza informações geradas pelo bate-papo on-line ao longo do tempo para extrair as principais dimensões latentes da satisfação do consumidor em relação à qualidade das marcas. De forma similar, Bendle e Wang (2016) apresentam uma técnica de análise de texto, também baseada no modelo Latent Dirichlet Allocation (LDA), que obtém percepções dos consumidores a partir das avaliações de produtos on-line.

Segundo He et al. (2015), análises competitivas de mídia social com referências de sentimento podem ser utilizadas para aprimorar a inteligência de marketing e melhorar a experiência dos clientes. Para isso, apresentam uma ferramenta denominada VOZIQ que analisa os tweets para gerar relatórios significativos de insight de negócios. De forma similar, Culotta e Cutler (2016) apresentam um método automatizado para classificação de percepções de marca através de atributos, explorando as conexões sociais das marcas no Twitter.

Para France e Ghose (2016), na era do big data e com a crescente disponibilidade de dados de compra do consumidor em larga escala, há a necessidade de técnicas que possam interpretar os dados e usá-los para tomadas de decisões gerenciais. Para isso, apresentam um método baseada na técnica de verossimilhança estatística para identificar, analisar e visualizar sub-mercados em categorias de produtos.

Chen, Zhang, Zhao (2017) apresentam uma abordagem analítica de *Big Data* que permite a identificação de clientes de alto valor, entendendo os padrões de uso e preferências no uso de aplicativos móveis, desenvolvida a partir do uso de três técnicas de mineração de dados: análise RFM (recência, frequência, monetária), análise de links e aprendizado de regras de associação. O objetivo do artigo apresentado por Verma e Singh (2017) é explorar várias limitações dos sistemas de mineração convencionais na extração de padrões de compra úteis de bancos de dados transacionais de varejo alimentados pelo Big Data. Os resultados da pesquisa revelam que os algoritmos de mineração convencionais não evoluíram para suportar a análise de *Big Data*, que exigem muitos recursos, como memória e mecanismos computacionais. Baseados nesses resultados, os autores desenvolvem o algoritmo MR-Apriori na forma de ferramenta IRM (*Infrastructure Resource Management*) para obter maior eficiência na mineração de dados.

Shimomura et al. (2018) propõem um novo método para identificar as orientações e os requisitos dos clientes para o desenvolvimento de projetos de sistemas de serviço-produto (PSS) empregando uma combinação de análise de tópicos, persona e cenários. Para isso, utilizam o

modelo *Latent Dirichlet Allocation* (LDA) no agrupamento de clientes para resolver a limitação de *clusters* de clientes baseados somente em dados demográficos e transacionais.

Por fim, Tian, Zhang e Zhang (2018) investigam como os parâmetros meteorológicos, incluindo a luz do sol, temperatura e qualidade do ar, podem ser incorporados aos sistemas de suporte a decisões de marketing (MDSS) para prever a variedade de consumidores em suas compras. As análises demonstram que as condições climáticas estão associadas a um maior comportamento de busca por variedade.

4.2.4 Aplicações do Big Data em Marketing

Os artigos desse grupo estão relacionados à aplicação do *Big Data* em estratégias, tomadas de decisões e ações na área de Marketing.

Wedel e Kannan (2016) avaliaram métodos de análise de marketing e aplicativos para dados estruturados e não estruturados, analisando seu potencial para apoiar decisões de marketing. Os autores identificam as direções para novos métodos de pesquisa analítica, abordando (1) análises para otimizar os gastos do mix de marketing em um ambiente rico em dados, (2) análises para personalização e (3) dados analíticos no contexto de privacidade e segurança de dados dos clientes. Eles analisaram as implicações para as organizações que pretendem implementar a análise de big data e tendências para a análises em marketing.

Com base em dados fornecidos pela National Hockey League (NHL) dos EUA, George e Wakefield (2018) estudaram o comportamento de compra dos consumidores e os contatos de vendas (comunicações de voz e texto do vendedor) para identificar fatores que atrapalham e impelem clientes a se associarem à instituição. Os resultados mostram que a quantidade, o tipo e o momento dos contatos de venda influenciam na efetivação da associação dos consumidores. Talón-Ballesteros et al. (2018) acessaram o perfil de clientes a partir de dados do sistema de CRM (*Customer Relationship Management*) de uma rede de hotéis internacional usando tecnologia de dados e técnicas de reamostragem *Bootstrap* para testes de proporção. O estudo contou com o perfil de 4.935.806 clientes diferentes no sistema de CRM durante os anos de 2013 e 2014 e o objetivo do trabalho era modelar o perfil dos clientes que visitavam a rede pela primeira vez e dos clientes chamados “repetidores”. Uma alta consistência foi encontrada no traço mais representativo que foi de “repetidores” viajando sem filhos. Para as redes de hotéis para férias, os resultados sugeriram uma análise mais aprofundada sobre a orientação do público-alvo para novos segmentos de mercado. O estudo mostra que a tecnologia de dados é especialmente útil para a análise interna de dados de CRM.

Anshari et al. (2015) investigaram o impacto do CRM social no sistema de aprendizagem online e o conceito de CRM social no contexto do *Big Data*. Eles conduziram uma pesquisa com 856 estudantes em Brunei Darussalam (Brunei) para explorar seus pontos de vista e a relação do aprendizado online usando o CRM social. A coleta de dados usou amostragem por conveniência. Os resultados do estudo mostram que o uso do CRM social no aprendizado online proporcionou uma mudança de paradigma de um meio de aprendizagem centrado e estático para um meio de aprendizagem compartilhado e dinâmico.

Os resultados de uma pesquisa preliminar sobre as atitudes dos consumidores a partir do uso de *Big Data* em Marketing realizada na Croácia são apresentados por Čurko, Merkaš e Silović (2018) em seu artigo que traz as características básicas e as aplicações do *Big Data* nos negócios. O estudo dá uma ênfase especial à aplicação do *Big Data* no Marketing com uma visão geral dos desafios fundamentais que a tecnologia de big data traz para as atividades de Marketing. O documento fornece recomendações sobre como responder a esses desafios.

Com base em uma abordagem de programação matemática, o artigo de Quin et al. (2017) estudou as estratégias de DSPs (*Demand Side Platforms* ou Plataformas pelo lado da Demanda) para segmentação de mercado e estabeleceu um modelo de seleção da granularidade para segmentar os mercados de publicidade de RTB (*Real Time Bidding* ou Concorrência em

Tempo Real). Com a abordagem do experimento computacional, projetaram três cenários experimentais para validar o modelo proposto e os resultados experimentais mostraram que: 1) a segmentação de mercado tem o potencial de melhorar a receita total de todos os anunciantes; 2) com o crescente refinamento da granularidade de segmentação do mercado, a receita total tem uma tendência de aumentar inicialmente e depois cair; 3) a granularidade ideal da segmentação de mercado será significativamente influenciada pelo número de anunciantes na DSP, mas menos influenciada pelo número de solicitações de anúncios. Segundo os autores, as descobertas mostram o papel crucial da segmentação de mercado no efeito de publicidade de RTB e indicam que as DSPs devem ajustar suas estratégias de segmentação de mercado de acordo com o número total de anunciantes.

Martin, Borah e Plamattier (2017) partem da premissa de vários textos e artigos de marketing que reconhecem que o uso de dados do consumidor aumenta o retorno das ações de Marketing. Por outro lado, quanto mais dados se obtém, mais suscetível fica a privacidade do consumidor e o mesmo reage negativamente a esse risco de interferência em sua privacidade. Dessa forma, em seu artigo, os autores objetivam melhorar a compreensão do efeito da vulnerabilidade dos dados do cliente no comportamento do mesmo e no desempenho da empresa, bem como os principais mecanismos de mediação e estratégias de mitigação dessas vulnerabilidades.

Jin *et al.* (2016) afirmam que até o advento do *Big Data* as empresas realizavam pesquisas com consumidores a partir de amostras limitadas para poder identificar os requisitos do consumidor para os produtos e assim trazer inovações e atualizações. Por outro lado, apontam que milhões de opiniões sobre os produtos, características e desempenhos estão disponíveis em websites como Amazon.com, mas muitos deles não são igualmente úteis, alguns por serem longos, outros por serem objetivos, mas, às vezes, fundamentais. De acordo com os autores isso é explicado pelos V's do *Big Data*. Os autores ponderam que embora os Vs superem na análise dos RC's (requisitos do consumidor ou *customer requests*) a capacidade de muitos modelos convencionais que foram construídos com base em dados limitados de pesquisa com clientes para *design* de produtos orientados para o mercado, no campo de gestão de clientes ou para a engenharia de *design*, algoritmos criados exclusivamente para ajudar os administradores de empresas, engenheiros de pesquisa e cientistas de dados a entenderem os RC's de maneira eficaz e eficiente a partir de grandes bases de dados de opinião ainda não estão disponíveis. Isso os fez identificar um espaço ainda não ocupado em pesquisa, com significância prática, que é desenvolver um modelo para lidar com os grandes dados de opinião do consumidor para o entendimento dos RC's. Assim, como base para o artigo, conduziram uma pesquisa para explorar o valor desses dados na perspectiva dos *designers* de produtos. Ela começa com a identificação de características do produto e das polaridades de sentimento a partir de grandes dados de opinião do consumidor. O método de filtro de Kalman é então empregado para prever as tendências de RC's e o método Bayesiano é proposto para comparar produtos. Por fim, usando dados de opinião colhidos na Amazon.com, um estudo de caso é apresentado para ilustrar como as técnicas propostas podem ser aplicadas.

Hofacker, Malthouse e Sultan (2016) desenvolveram um documento que busca avaliar como o estudo do comportamento do consumidor pode ser beneficiar da análise de big data. O artigo oferece uma visão geral conceitual das oportunidades potenciais e mudanças no estudo do comportamento do consumidor que o *Big Data* provavelmente trará. A perspectiva dos autores é que o *Big Data* tem o potencial de ampliar a compreensão de cada etapa do processo de tomada de decisão do consumidor. Embora o estudo do comportamento do consumidor tenha tradicionalmente avançado usando uma teoria seguida de experimentação, agora parece que a natureza do ciclo de retroalimentação entre a teoria e os resultados pode mudar sob o peso do *Big Data*.

A análise do consumidor está no epicentro de uma revolução do *Big Data*. A tecnologia ajuda a capturar dados ricos e abundantes sobre os fenômenos do consumidor em tempo real. Assim,

volume, velocidade e variedade de dados primários sem precedentes (*Big Data*) estão disponíveis para consumidores individuais. A partir dessa observação, Erevelles *et al.* (2016), propõem uma estrutura conceitual que se apoia na teoria baseada em recursos com o propósito de entender melhor o impacto do *Big Data* em várias atividades de marketing, permitindo que as empresas explorem melhor seus benefícios.

Dinu *et al.* (2016) em seu artigo discorrem sobre o fato de que com o advento do desenvolvimento da internet e da tecnologia, ocorreram mudanças no comportamento de consumo das pessoas. Até recentemente, o consumidor não tinha acesso a uma riqueza de informações sobre um determinado produto, nem uma ampla gama de produtos disponíveis para escolher no momento da compra. Mas agora, em um século de velocidade e mudança permanente, cada vez mais consumidores recorrem à Internet para obter informações, fazer compras ou realizar transações bancárias. Essas atividades on-line são registradas como conjuntos de dados e formam o *Big Data*. A tecnologia *Big Data* permite que os consumidores encontrem o produto certo ou as informações desejadas, e as empresas compreendam o comportamento de seus clientes.

Com a proposta de que existe uma relação complementar entre a natureza formalizada dos dados do cartão de fidelidade digital e a natureza informal da orientação para o mercado de pequenas empresas, Donnelly *et al.* (2015) analisaram essa relação em pequenas empresas que tiveram acesso aos dados do cartão do Clube Tesco, usando uma abordagem de pesquisa baseada em caso. As evidências revelaram uma nova estrutura e precisão no planejamento de marketing das pequenas empresas que tiveram acesso a esses dados.

A evolução dos gostos pode mudar a importância relativa das características do produto na definição das decisões de compra dos consumidores, o que, por sua vez, pode alterar a relativa atratividade de produtos com diferentes níveis de recursos. (Du, Hu e Damangi, 2015). O desafio, para os autores, está em encontrar uma maneira confiável e econômica de monitorar os pesos que os consumidores atribuem às várias características do produto. O artigo explora o potencial de usar tendências em pesquisas de palavras-chave on-line como indicadores de tendências na importância relativa dos benefícios correspondentes (por exemplo, economia de combustível, aceleração, custo para comprar, custo para operar, tipo de corpo) a partir do contexto da indústria automobilística americana. Ao aumentar os dados do composto mercadológico com dados das buscas por características em um modelo de resposta do mercado, eles mostram melhorias substanciais na qualidade de adequação dentro e fora da amostra.

O desenvolvimento da internet trouxe a criação de comunidades de vários tipos, entre elas comunidades especializadas em temas, nas quais seus membros colaboram entre si, compartilhando dúvidas e conhecimento. O artigo de Bone *et al.* (2015) destaca que as empresas *business to business* estão também, cada vez mais, criando comunidades de solução de problemas pessoa a pessoa (P3) suportadas pela empresa, para atender algumas de suas necessidades de serviço de suporte ao cliente, uma vez que uma questão importante que essas empresas enfrentam é como reduzir a dependência dos clientes do serviço de assistência tradicional. Com isso, os autores levantam a questão: a participação na solução de problemas do cliente nessas comunidades reduz as demandas associadas ao serviço tradicional de atendimento ao cliente? Baseados nessa pergunta, realizaram uma pesquisa para investigar os efeitos da participação de clientes com problemas a resolver em uma comunidade P3 entre os clientes *business to business* globais. Os resultados da investigação revelaram que a participação do cliente com problemas a resolver na comunidade, em termos de ajudar a si mesmo (postar perguntas) e ajudar aos outros (respondendo a perguntas de colegas), reduz o uso do serviço de suporte ao cliente tradicional.

Stevens (2014) explora em seu artigo a questão de que o setor cultural precisa de informações para apoiar seus esforços de recapitalização. Contudo, a grande questão é como avaliar a

quantidade de capital que pode ser levantada. Segundo o autor, através da big data, agora é possível entender o universo do apoio ao setor cultural. Em um teste que empregou mais de dois milhões de células de dados de 2007 em diante, das organizações culturais em Charlotte, Carolina do Norte, o estudo encontrou padrões e prognósticos evidentes de apoio ao setor como base para recapitalização, demonstrando a relevância do big data.

O contínuo crescimento da internet globalmente chama muita atenção para as imensas oportunidades que o marketing digital apresenta, com pouca atenção para os reais desafios que as empresas enfrentam para se tornarem digitais (Leeflang *et al.*, 2014). Em seu estudo, apresentam esses desafios com base nos resultados de uma pesquisa entre uma amostra por conveniência de 777 executivos de marketing em todo o mundo.

Rust e Huang (2014) afirmam que a natureza da ciência do marketing está mudando de maneira sistemática, previsível e irrevogável. Como a tecnologia da informação permite a comunicação onipresente com o cliente e muitos dados sobre o cliente, a natureza fundamental da conexão da empresa com o cliente muda. Essas mudanças implicam em uma transformação da ciência do marketing tanto nos tópicos a serem enfatizados quanto nos métodos a serem empregados. Cada vez mais, e inevitavelmente, todo o marketing se assemelhará em maior grau à área anteriormente especializada em marketing de serviços, apenas com uma ênfase crescente na parte analítica de marketing.

Weiberg *et al.* (2013) destacam em seu artigo que há muita publicidade associada ao termo “*Big Data*”, e muita oportunidade nos dados que estão associados a esse termo, juntamente com as ferramentas e técnicas existentes e ainda sendo desenvolvidas para alavancá-lo para a tomada de decisões e melhorar a condição dos seres vivos, empresas e sociedade. O foco de sua pesquisa foi explorar o significado de big data e identificar caminhos importantes de pesquisa sobre *Big Data*. Como parte desse processo, recorreram a um conjunto diversificado de especialistas em marketing que possuem conhecimentos e pontos de vista especiais. Assim, descobriram que diferentes comunidades têm perspectivas diferentes.

4.2.5 Pesquisa de Big Data em Marketing

Nesse grupo estão apresentados os artigos que discutem a evolução da pesquisa sobre big data na área de Marketing.

Donilcar e Ring (2014) realizaram uma revisão da pesquisa de marketing em turismo, concluindo que os trabalhos nessa área focam principalmente na forma como promessas de serviço são feitas e mantidas, em estruturas para tomada de decisões gerenciais ou insights sobre associações entre constructos, mas que estudos de relações de causa e efeito, são raros. Apontam então para oportunidades de pesquisas futuras, onde afirmam que o Big Data tem o potencial de gerar grandes quantidades de conhecimento de alta qualidade.

Em uma pesquisa bibliográfica baseada em *text mining*, utilizando publicações de 2010 a 2015, Amado *et al.* (2018) analisaram as principais tendências do Big Data na área de marketing. Para eles, a pesquisa nesse tema se apresenta em estágio embrionário, bipartida entre os domínios tecnológico e de pesquisa, com publicações de Big Data não se alinhando claramente as técnicas de ponta com os benefícios do Marketing.

Berrio, R. *et al.* (2016) apresentam o big data como uma tendência emergente em pesquisa de mercado. Os autores realizaram uma revisão sistemática da literatura e usaram para isso duas fontes secundárias: o Instituto de Ciência do Marketing (*Marketing Science Institute*) e cinco equações de busca no Scopus. A equação de busca que foi selecionada gerou 243 resumos no Scopus para a janela de observação de 2005 a 2015, mostrando que o Big Data tem relevância na dinâmica internacional do conhecimento e constitui uma ferramenta útil para a gestão de mercado, sendo uma estratégia de gerenciamento de informações.

5 Considerações finais

Com base na amostra de artigos obtida, os resultados indicam que a maioria dos estudos de Big Data na área de marketing focam em métodos, técnicas e uso de ferramentas de mineração e análise de dados originários de banco de dados formados por registros de transações comerciais e avaliações realizadas por consumidores e clientes. As pesquisas são baseadas tanto em dados estruturados quanto não-estruturados, ou seja, enquanto algumas pesquisas utilizam dados estruturados, minerados em bancos de dados empresariais, outros recorrem a dados não estruturados, obtidos em bases de dados originários de sites e redes sociais. Em menor número, foram encontradas pesquisas realizadas com técnicas de coleta e análise de Big Data e análises de publicações e pesquisas sobre Big Data em marketing realizadas anteriormente.

As principais limitações da pesquisa estão relacionadas à escolha de uma única base de dados (Scopus). Apesar de oferecer melhor cobertura na área de ciências sociais (Norris e Oppenheim, 2007), maior número de publicações indexadas (Bar-Ilan, 2008) e maior número de citações (Meho e Yang, 2007) em relação à base Web of Science (WOS), podem haver artigos únicos nas diferentes bases de dados, conforme o discutido por Meho e Yang (2007).

Como sugestão para novas pesquisas destaca-se a necessidade de novas pesquisas envolvendo não somente a discussão das técnicas de mining e análise dos dados, mas também estudos que envolvam a utilização prática das análises de Big Data e respectivos resultados em estratégias e ações de marketing.

6 Referências

- 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.
- Anshari, M., Alas, Y., Yunus, N., Sabtu, N. I., & Hamid, M. H. (2015). Social customer relationship management and student empowerment in online learning systems. *International Journal of Electronic Customer Relationship Management*, 9(2-3), 104-121.
- Banerjee, S., Viswanathan, V., Raman, K., & Ying, H. (2013). Assessing prime-time for geotargeting with mobile big data. *Journal of Marketing Analytics*, 1(3), 174-183.
- Bar-Ilan, J. (2008). Which h-index? - A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, 74(2), 257-271.
- Bardin, L. (2004). *Análise de conteúdo*. 3ª. Lisboa: Edições, 70.
- Berrio, S. P. R., Rodriguez, R. A. V., & ROBAYO, Ó. J. (2016). Big Data: trend emerging from research in marketing. *Espacios*, 37(38).
- Bone, S. A., Fombelle, P. W., Ray, K. R., & Lemon, K. N. (2015). How customer participation in B2B peer-to-peer problem-solving communities influences the need for traditional customer service. *Journal of Service Research*, 18(1), 23-38.
- Bornmann, L., & Daniel, H. D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45-80.
- Canaltech. O que é Big Data? (2018, september 13). São Paulo, 2018. Retrieved from: <<https://canaltech.com.br/big-data/o-que-e-big-data/>> Acesso em:
- Chiang, L. L. L., & Yang, C. S. (2018). Does country-of-origin brand personality generate retail customer lifetime value? A Big Data analytics approach. *Technological Forecasting and Social Change*, 130, 177-187.
- Chen, Q., Zhang, M., & Zhao, X. (2017). Analysing customer behaviour in mobile app usage. *Industrial Management & Data Systems*, 117(2), 425-438.
- Cronin, B. (2001). Bibliometrics and beyond: some thoughts on web-based citation analysis. *Journal of Information science*, 27(1), 1-7.
- Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field. *Journal of Informetrics*, 5(1), 146-166.

- Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2012). SciMAT: A new science mapping analysis software tool. *Journal of the American Society for Information Science and Technology*, 63(8), 1609-1630.
- Cox, M., & Ellsworth, D. (1997). Application-controlled demand paging for out-of-core visualization. In *Visualization'97., Proceedings* (pp. 235-244). IEEE.
- Ćurko, K., Merkaš, Z., & Silović, T.. (2018). Challenges of application of the big data in marketing: Case study Croatia. *World Scientific and Engineering Academy and Society*
- Dinu, D., Stoica, I., & RADU, A. V. (2016). Studying the Consumer Behavior through Big Data. *Quality-Access to Success*, 17.
- Dolnicar, S., & Ring, A. (2014). Tourism marketing research: Past, present and future. *Annals of Tourism Research*, 47, 31-47.
- Donnelly, C., Simmons, G., Armstrong, G., & Fearn, A. (2015). Digital loyalty card 'big data' and small business marketing: Formal versus informal or complementary?. *International Small Business Journal*, 33(4), 422-442.
- Du, R. Y., Hu, Y., & Damangir, S. (2015). Leveraging trends in online searches for product features in market response modeling. *Journal of Marketing*, 79(1), 29-43.
- Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897-904.
- France, S. L., & Ghose, S. (2016). An analysis and visualization methodology for identifying and testing market structure. *Marketing Science*, 35(1), 182-197.
- Francisco, Eduardo (2014). Big Data ou Big problems? *Revista da ESPM Ano 20, Edição 95, Nº 5*.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.
- He, W., Wu, H., Yan, G., Akula, V., & Shen, J. (2015). A novel social media competitive analytics framework with sentiment benchmarks. *Information & Management*, 52(7), 801-812.
- Hyrnsalmi, S., Seppanen, M., Aarikka-Stenroos, L., Suominen, A., Jarvelainen, J., & Harkke, V. (2015). Busting myths of electronic word of mouth: the relationship between customer ratings and the sales of mobile applications. *Journal of theoretical and applied electronic commerce research*, 10(2), 1-18.
- Hofacker, C. F., Malthouse, E. C., & Sultan, F. (2016). Big data and consumer behavior: Imminent opportunities. *Journal of Consumer Marketing*, 33(2), 89-97.
- Irudeen, R., & Samaraweera, S. (2013, December). Big data solution for Sri Lankan development: A case study from travel and tourism. In *Advances in ICT for Emerging Regions (ICTer), 2013 International Conference on* (pp. 207-216). IEEE.
- Jin, J., Liu, Y., Ji, P., & Liu, H. (2016). Understanding big consumer opinion data for market-driven product design. *International Journal of Production Research*, 54(10), 3019-3041.
- Jun, S. P., Park, D. H., & Yeom, J. (2014). The possibility of using search traffic information to explore consumer product attitudes and forecast consumer preference. *Technological Forecasting and Social Change*, 86, 237-253.
- Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META group research note*, 6(70), 1.
- Leeftang, P. S., Verhoef, P. C., Dahlström, P., & Freundt, T. (2014). Challenges and solutions for marketing in a digital era. *European management journal*, 32(1), 1-12.
- Liu, Y. P., Guo, J. F., & Fan, Y. (2017). A big data study on emitting companies' performance in the first two phases of the European Union Emission Trading Scheme. *Journal of Cleaner Production*, 142, 1028-1043.
- Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data privacy: Effects on customer and firm performance. *Journal of Marketing*, 81(1), 36-58.

- Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the american society for information science and technology*, 58(13), 2105-2125.
- Moya-Anegón, F., Chinchilla-Rodríguez, Z., Vargas-Quesada, B., Corera-Álvarez, E., Muñoz-Fernández, F. J., González-Molina, A., & Herrero-Solana, V. (2007). Coverage analysis of Scopus: A journal metric approach. *Scientometrics*, 73(1), 53-78.
- Norris, M., & Oppenheim, C. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences' literature. *Journal of informetrics*, 1(2), 161-169.
- Narin, F., & Hamilton, K. S. (1996). Bibliometric performance measures. *Scientometrics*, 36(3), 293-310.
- Qin, R., Yuan, Y., & Wang, F. Y. (2017). Exploring the optimal granularity for market segmentation in RTB advertising via computational experiment approach. *Electronic Commerce Research and Applications*, 24, 68-83.
- Rust, R. T., & Huang, M. H. (2014). The service revolution and the transformation of marketing science. *Marketing Science*, 33(2), 206-221.
- SAS. O que é Big Data? (2018, September 13). Retrieved from: <https://www.sas.com/pt_br/insights/big-data/what-is-big-data.html>.
- SCImago, (2018). SJR — SCImago Journal & Country Rank [Portal]. Retrieved 30/10/2018, from <http://www.scimagojr.com>
- Shimomura, Y., Nemoto, Y., Ishii, T., & Nakamura, T. (2018). A method for identifying customer orientations and requirements for product-service systems design. *International Journal of Production Research*, 56(7), 2585-2595.
- Stevens, L. K. (2014). Making the Case: Using a Living and Historic Cultural Universe Big Data System as a Knowledge-Based Foundation for Sustaining and Recapitalizing the Cultural Sector. *The Journal of Arts Management, Law, and Society*, 44(4), 208-217.
- Talón-Ballester, P., González-Serrano, L., Soguero-Ruiz, C., Muñoz-Romero, S., & Rojo-Álvarez, J. L. (2018). Using big data from Customer Relationship Management information systems to determine the client profile in the hotel sector. *Tourism Management*, 68, 187-197.
- Tian, J., Zhang, Y., & Zhang, C. (2018). Predicting consumer variety-seeking through weather data analytics. *Electronic Commerce Research and Applications*, 28, 194-207.
- Tirunillai, S., & Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. *Journal of Marketing Research*, 51(4), 463-479.
- Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538.
- Van Raan, A. F. (2004). Measuring science. In *Handbook of quantitative science and technology research* (pp. 19-50). Springer Netherlands.
- Verma, N., & Singh, J. (2017). An intelligent approach to Big Data analytics for sustainable retail environment using Apriori-MapReduce framework. *Industrial Management & Data Systems*, 117(7), 1503-1520.
- Wallin, J. A. (2005). Bibliometric methods: pitfalls and possibilities. *Basic & clinical pharmacology & toxicology*, 97(5), 261-275.
- Weinberg, B. D., Davis, L., & Berger, P. D. (2013). Perspectives on big data. *Journal of Marketing Analytics*, 1(4), 187-201.