

**TURISMO E BIG DATA NAS PESQUISAS CIENTÍFICAS BRASILEIRAS**

TOURISM AND BIG DATA AT BRAZILIAN SCIENTIFIC RESEARCH

TURISMO Y BIG DATA EN LA INVESTIGACIÓN CIENTÍFICA BRASILEÑA

Natalya Reis da Silva^{1*} & **Erika Sayuri Koga**²^{1,2} Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, Campus São Paulo.¹ natalya.reis@aluno.ifsp.edu.br ² kogadinapoli@ifsp.edu.br**RESUMO INFO.****Recebido:** 18.08.2022**Aprovado:** 22.08.2022**Disponibilizado:** 23.08.2022**PALAVRAS-CHAVE:** Big data; Turismo; Modelo interdisciplinar; Revisão sistemática da literatura.**KEYWORDS:** Big data, Tourism; Interdisciplinary model; Systematic Review of Literature.**PALABRAS CLAVE:** Big data; Turismo; Modelo interdisciplinario; Revisión sistemático de literatura***Autor Correspondente:** Silva, N. R.**RESUMO**

O aumento da capacidade de armazenamento e velocidade de processamento dos computadores, aliadas às rápidas conexões da internet, desenvolvimento de algoritmos e disponibilidade de grande volume de dados na rede são fatores que impulsionam a era da *big data*. Considerando o crescimento e a diversidade de possibilidades do uso de *big data* para pesquisas no turismo realizadas mundialmente, esta investigação se propõe a realizar uma revisão sistemática sobre uso de *big data* nas pesquisas de turismo, conduzidas por pesquisadores brasileiros e publicadas nas revistas nacionais e internacionais. Após filtrar 76 artigos extraídos nas bases de dados *Web of Science*, *Scopus* e Publicações de Turismo, esta revisão analisou 19 artigos que atenderam aos critérios estabelecidos no protocolo. Verificaram-se que os estudos com o uso de *big data* no turismo são recentes, a partir de 2000, com um pico de estudos publicados em 2019; realizados por uma variedade de pesquisadores (57), alocados em diferentes Instituições de Pesquisa/Ensino (16), distribuídos pelo Brasil, Portugal e Espanha, que publicaram em diferentes revistas científicas nacionais e internacionais (15). As pesquisas analisadas contribuem com vários campos do conhecimento do turismo, conforme Modelo Interdisciplinar de Jafari, sobretudo para o gerenciamento de empresas, planejamento e desenvolvimento do turismo e motivação dos turistas, demonstrando um nicho de pesquisa ainda em expansão e com poucos destaques. Os resultados evidenciaram, também, que 47,4% dos artigos têm uma abordagem metodológica qualitativa, 63,2% realizaram pesquisas cujas fontes dos dados foram geradas por usuários e que as universidades que apresentam a maior quantidade de publicações estão localizadas na região nordeste do Brasil e são federais. Ademais, este estudo pode contribuir para o reconhecimento da produção científica dos brasileiros que desenvolveram pesquisas com o uso de *big data* e fomentar intercâmbio de técnicas, softwares, processos de coleta, análises e obtenção de resultados que fortaleçam os estudos no turismo e superem desafios tanto de ordem procedural e conceitual, quanto de ordem teórica e metodológica.

ABSTRACT

The increase in storage capacity and computer processing speed, combined with fast internet connections, development of algorithms and availability of large volumes of data on the network are factors that drive the era of *big data*. Considering the growth and diversity of possibilities for the use of *big data* for tourism research carried out worldwide, this investigation proposes to carry out a systematic review on the use of *big data* in tourism research, conducted by Brazilian researchers and published in national and international journals. After filtering 76 articles extracted from the Web of

Science, Scopus and Tourism articles databases, this review analyzed 19 articles that met the criteria established in the protocol. It was found that studies on the use of *big data* in tourism are recent, starting in 2000, with a peak of studies published in 2019; carried out by a variety of researchers (57), allocated in different Research/Education Institutions (16), distributed across Brazil, Portugal and Spain, who published in different national and international scientific journals (15). The analyzed research contributes to several fields of tourism knowledge, according to Jafari's Interdisciplinary Model, especially for business management, tourism planning and development and tourist motivation, demonstrating a research niche that is still expanding and with few highlights. The results also showed that 47.4% of the articles have a qualitative-quantitative methodological approach, 63.2% carried out research whose data sources were generated by users and that the universities with the largest number of publications are in the region northeast of Brazil and are federal. Furthermore, this study can contribute to the recognition of the scientific production of Brazilians who have developed research using *big data* and foster the exchange of techniques, software, collection processes, analyzes and obtaining results that strengthen studies in tourism and overcome challenges both procedural and conceptual, as well as theoretical and methodological.

RESUMEN

El aumento de la capacidad de almacenamiento y la velocidad de procesamiento de las computadoras, combinado con conexiones rápidas a Internet, el desarrollo de algoritmos y la disponibilidad de grandes volúmenes de datos en la red son factores que impulsan la era del *big data*. Considerando el crecimiento y la diversidad de posibilidades para el uso de *big data* para investigaciones turísticas realizadas en todo el mundo, esta investigación propone realizar una revisión sistemática sobre el uso de *big data* en investigaciones turísticas, realizada por investigadores brasileños y publicada en revistas nacionales e internacionales. Tras filtrar 76 artículos extraídos de las bases de datos *Web of Science*, *Scopus* y publicaciones de Turismo, esta revisión analizó 19 artículos que cumplían con los criterios establecidos en el protocolo. Se encontró que los estudios sobre el uso de *big data* en turismo son recientes, inician en el año 2000, con un pico de estudios publicados en el 2019; realizado por una variedad de investigadores (57), asignados en diferentes Instituciones de Investigación/Educación (16), distribuidos en Brasil, Portugal y España, que publicaron en diferentes revistas científicas nacionales e internacionales (15). Las investigaciones analizadas contribuyen a varios campos del conocimiento turístico, según el Modelo Interdisciplinario de Jafari, especialmente para la gestión empresarial, la planificación y el desarrollo turístico y la motivación turística, demostrando un nicho de investigación aún en expansión y con pocos destaques. Los resultados también mostraron que el 47,4% de los artículos tienen un enfoque metodológico cualitativo-cuantitativo, el 63,2% realizó investigaciones cuyas fuentes de datos fueron generadas por los usuarios y que las universidades con mayor número de publicaciones están ubicadas en la región noreste de Brasil, y son federales. Además, este estudio puede contribuir al reconocimiento de la producción científica de los brasileños que han desarrollado investigaciones utilizando *big data* y fomentar el intercambio de técnicas, software, procesos de recolección, análisis y obtención de resultados que fortalecen los estudios en turismo y superan desafíos tanto procedimentales como conceptuales, así como teóricos y metodológicos.



INTRODUÇÃO

O crescente poder dos computadores providos pelo aumento de suas capacidades de armazenamento e velocidade de processamento, aliada às conexões de internet cada vez mais rápidas, ao desenvolvimento de algoritmos poderosos e à disponibilidade de grandes quantidades de dados são alguns dos fatores que vêm impulsionando o desenvolvimento de análises, com possibilidade de contribuir com o conhecimento da ciência, aprimoramento dos negócios e também para a sociedade em geral (Egger, 2022). Acompanhar a evolução dessa nova era do *big data* em diversas áreas, abre-se o olhar para suas aplicações e potencialidades no turismo (Li *et al.*, 2018), com possibilidades de realizar estudos orientados por dados coletados em diversas fontes, cujas origens são caracterizadas pelo *big data*.

A expansão de ferramentas e métodos de pesquisa ocorrendo, sobretudo, pela Ciência de Dados, abre oportunidades para as pesquisas no campo do turismo. O desafio que se coloca também é sobre os métodos científicos, técnicas e instrumentos a serem utilizados em pesquisas de análise de *big data* que resolvam problemas de pesquisas, cumpram seus objetivos propostos e tragam contribuições para o referencial teórico em construção.

Imagine a facilidade para agências, hotéis ou outros empreendimentos turísticos fazer recomendações customizadas aos turistas de acordo com seus desejos, necessidades e interesses, cujos dados foram coletados através de ferramentas de Inteligência Artificial (IA) e processados por *Machine Learning* (ML) e *Deep Learning* (DL) (Egger, 2022). Com o crescimento dos comentários *online* deixados por hóspedes sobre suas experiências nos hotéis, pesquisas têm sido realizadas para converter em treinamentos e melhorias na qualidade de serviços e estruturas. Zheng Xiang *et al.* (2015) realizaram estudos a partir da análise de *big data* extraídos do site Expedia.com e identificaram que há forte relação entre os relatos de experiências dos usuários e *ranking* de satisfação dos hóspedes. Dados extraídos de blogs de viagem também podem ser utilizados para coleta de dados sobre análise de sentimentos dos turistas e recomendações de destinos e roteiros de viagens. Hua Yuan *et al.* (2016) realizam análise de diversos blogs de viagens sobre uma cidade chinesa específica para definir um método padrão de mineração de dados que possibilite identificar atividades turísticas mais usuais.

OBJETIVO

O objetivo desta pesquisa foi realizar uma revisão sistemática sobre *big data* nas pesquisas de turismo conduzidas por pesquisadores brasileiros e publicadas nas revistas nacionais e internacionais. Complementarmente, conduziu-se as análises para: identificar os principais centros de pesquisa ou universidades no Brasil que estão produzindo pesquisas com uso de *big data*; listar as revistas científicas e os pesquisadores com publicações sobre o tema; identificar as contribuições com os campos de conhecimento do turismo; e, acompanhar a evolução dessas publicações por ano.

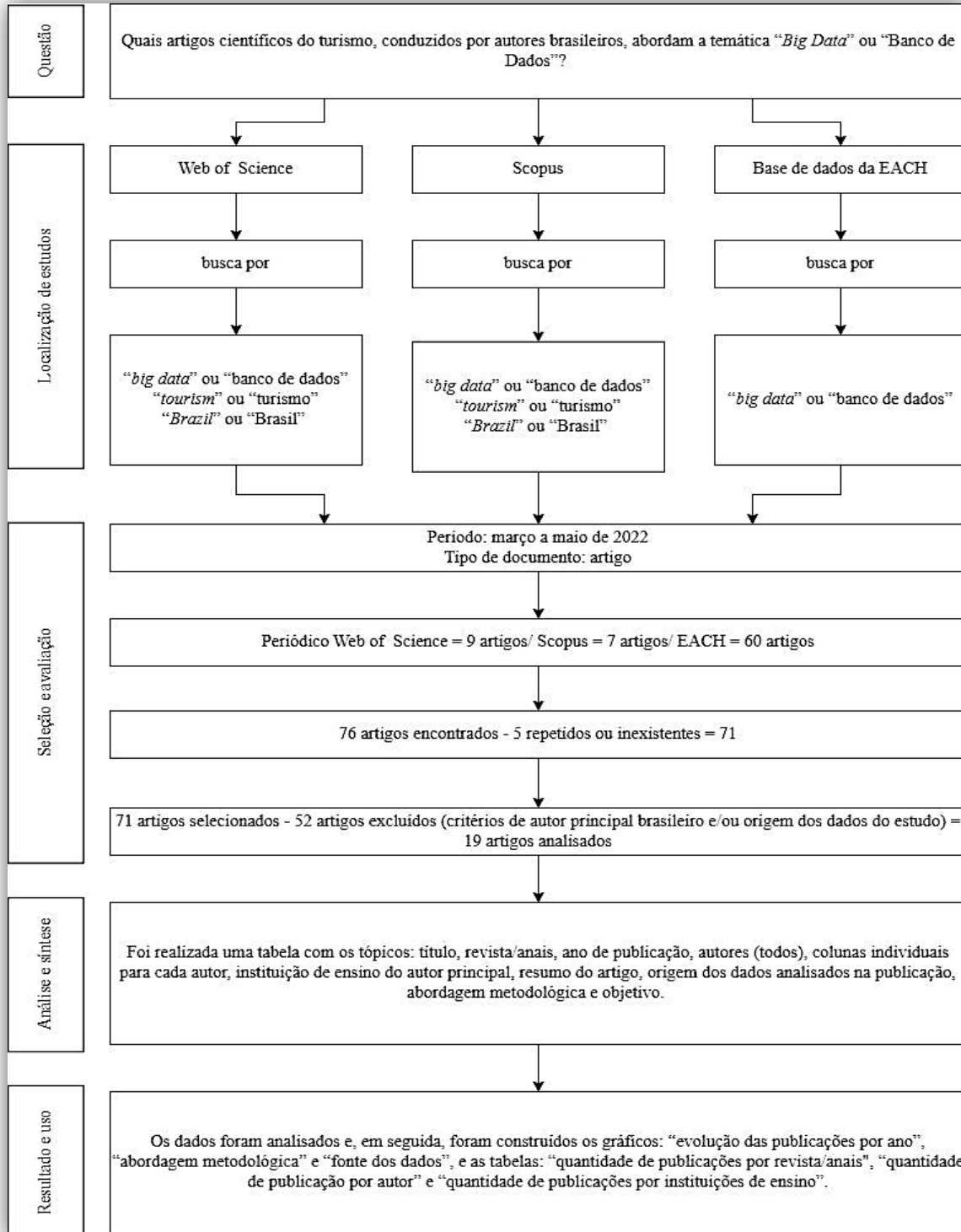
MATERIAL E MÉTODO

Esta pesquisa de caráter exploratório conduziu uma revisão sistemática da literatura (Atallah & Castro, 1998) de artigos publicados por brasileiros nos âmbitos nacional e internacional, realizando buscas por artigos publicados em revistas científicas e anais de eventos indexados nas bases de dados *Web of Science*, *Scopus* e Publicações de Turismo. A revisão sistemática



seguir um protocolo desenvolvido especificamente para alcançar os objetivos propostos deste artigo para documentar os procedimentos de análises e definição de critérios (Rother, 2007). A partir das buscas por palavras-chave e inclusão dos critérios estabelecidos para pesquisa, as análises focaram em 19 artigos científicos. O processo da revisão sistemática ocorreu conforme mostra a figura 1.

Figura 1. Fluxograma da pesquisa



Fonte: Autores (2022)



RESULTADOS E DISCUSSÃO

Em síntese, os resultados da revisão sistemática indicam que o ano de 2019 teve o maior número de publicações de artigos dentro dos critérios propostos neste estudo, mas não foi possível identificar nenhuma tendência de crescimento ou decréscimo. A análise evidencia, também, que 47,4% dos artigos têm uma abordagem metodológica quali-quantitativa, seguida de forma igualitária de pesquisas com abordagens quantitativas (26,3%) e qualitativas (26,3%). Quanto às fontes de dados utilizadas nas pesquisas, 63,2% realizaram pesquisas com dados gerados por usuários, em seguida, usaram-se dados coletados de dispositivos (21,1%) e, em terceiro, dados extraídos de operações ou transações (15,8%). Dentre os periódicos listados, a Revista Brasileira de Ecoturismo possui um breve destaque com três publicações, e segue com a Revista Turismo em Análise e a Licere, com duas publicações cada. Quanto aos autores, os resultados mostram que 53 autores publicaram apenas um artigo nessa área e apenas 4 pesquisadores têm duas publicações, de modo que dois autores são da Universidade Federal do Vale do São Francisco, e os outros dois autores da Universidade Federal de Sergipe, é interessante notar que os principais centros de pesquisa são Universidades Federais e estão localizadas na região nordeste do Brasil. A área do conhecimento com maior contribuição foi “gerenciamento de empresas” com 26,3% dos artigos analisados, em seguida “planejamento e desenvolvimento do turismo” e “motivação do turista” com o mesmo percentual (15,8%) e na sequência “Projeto com natureza”, “Legislação” e “Geografia do turismo” (10,5%).

CONCLUSÃO

Contribuir com a evolução dos métodos e técnicas de pesquisas científicas requer um começo. O início do processo vem com muita curiosidade e a evolução se dá com muito esforço, dedicação e estudo. Esses foram os itens motivadores desta revisão sistemática da literatura focada no uso de *big data* nas pesquisas de turismo realizadas por pesquisadores brasileiros. Este foco se justifica na intenção de promover trocas e diálogos mais próximos a quem já vem desenvolvendo pesquisas com o uso de *big data* no Brasil. Identificar onde estão, quem são, o que já fizeram, com quais objetivos, usos de quais fontes de dados e tipos de dados... são descobertas importantes para uma jornada científica promissora.

Assim, muitos caminhos se descortinam a partir dessa revisão de literatura, sobretudo em aprofundar nas análises dos procedimentos metodológicos realizados e os objetivos de pesquisas estabelecidos, entendendo que a evolução do campo de conhecimento do turismo exige essa rigorosidade na relação objetivo, metodologia com embasamento teórico-conceitual, contribuindo para o fortalecimento do turismo como uma ciência (Lohmann & Panosso Netto, 2008) e também favorecendo a evolução concomitante dos usos de dados coletados e construção das teorias (Mazanec, 2020).

Inspirada na revisão sistemática da literatura realizada por Li et al. (2018), esta pesquisa dará continuidade no aprofundamento das análises de formas utilizadas para coletas de dados, seguindo para o tratamento e organização das análises e obtenção de resultados dos estudos realizados pelos brasileiros, no intuito de identificar melhor as vantagens e desvantagens dos diversos métodos existentes e em uso no Brasil. Apesar da evidência identificada com essa revisão sistemática de que as pesquisas com *big data* no turismo ainda estão em fase embrionária, acredita-se na potencialidade e desdobramentos importantes para contribuir com o avanço das pesquisas no turismo.



REFERÊNCIAS

Atallah, A. N. & Castro, A. A. (1998). Revisão sistemática da literatura e metanálise. *Medicina baseada em evidências: fundamentos da pesquisa clínica*. São Paulo: Lemos-Editorial, 42-48.

Hua, Y., Hualin, X., Yu, Q., & Yan, L. (2016). Make your travel smarter: Summarizing urban tourism information from massive blog data. *International Journal of Information Management*, 36(6B), 1306-1319, <https://doi.org/10.1016/j.ijinfomgt.2016.02.009>

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, 68, 301–323. <https://doi.org/10.1016/J.TOURMAN.2018.03.009>

Lohmann, G. & Panosso, A., Netto. (2008). *Teoria do Turismo: conceitos, modelos e sistemas*. São Paulo: Aleph.

Mazanec, J. A. (2020). Hidden theorizing in big data analytics: With a reference to tourism design research. *Annals of Tourism Research*, 83(May), 102931. <https://doi.org/10.1016/j.annals.2020.102931>

Mongeon, P. & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, 106(1), 213–228. <https://doi.org/10.1007/s11192-015-1765-5>

Rother, E. T. (2007). Revisão sistemática X revisão narrativa. *Acta paulista de enfermagem*, 20(2), v-vi.

Zheng Xiang, Zvi Schwartz, John H. Gerdes, Muzaffer Uysal, (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, 44, 120-130, <https://doi.org/10.1016/j.ijhm.2014.10.013>

