

Review

Not peer-reviewed version

Systematic Review of Big Data Applications in Decoding Consumer Behaviors

[Qiankun Liu](#) , [Rui Gang Wang](#) , [Muhabaiti Paretı](#) * , [Alessandra Castellini](#) , [Maurizio Canavari](#) †

Posted Date: 21 May 2025

doi: [10.20944/preprints202505.1668.v1](https://doi.org/10.20944/preprints202505.1668.v1)

Keywords: big data; consumer behavior; predictive analytics; machine learning; marketing research; systematic review; text mining; data quality



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

Systematic Review of Big Data Applications in Decoding Consumer Behaviors

Qiankun Liu ¹, Ruigang Wang ², Muhabaiti Pareti ^{1,3,*}, Alessandra Castellini ¹ and Maurizio Canavari ^{1,†}

¹ Department of Agricultural and Food Sciences, Alma Mater Studiorum-University of Bologna, Viale Giuseppe Fanin 50, 40127, Bologna, Italy

² Institute of Food and Nutrition Development, Ministry of Agriculture and Rural Affairs, 100081, Beijing, China

³ College of Economics and Management, Xinjiang Agricultural University, Nongda East Road 311, Urumqi, 830052, China

* Correspondence: muhabaiti86@xjau.edu.cn

† Prof. Maurizio Canavari has passed away.

Abstract: Big data is pivotal in understanding consumer behavior and predicting consumer decisions. However, research has predominantly focused on specific consumption aspects, with a noticeable gap in systematic reviews on big data's role in consumer behavior studies. This paper systematically reviews 127 articles to identify key topics, significance, challenges, and emerging trends in the application of big data to consumer behavior research. Our findings indicate that big data analysis in this field primarily focuses on consumer attitudes, behavior patterns, decision-making processes, and the impact of major events. Big data is categorized into structured and unstructured types, with deep learning, machine learning, and text data as essential research methods, particularly for predicting consumer trends. Future research should focus on enhancing data quality, improving model interpretability, and fostering stronger collaboration between academia and industry. This study advances the understanding of how big data can be effectively leveraged in consumer behavior research, highlighting its potential benefits and challenges.

Keywords: big data; consumer behavior; predictive analytics; machine learning; marketing research; systematic review; text mining; data quality

1. Introduction

Recently, big data analysis and consumer behavior have played an important role in transforming the entire marketing process and how companies and market researchers analyze consumer behavior and market trends. As a result, it has been receiving significant focus from researchers in consumer behavior and marketing (L. Li, 2023). Researchers have been using different types of big data and various sophisticated methods, models, and algorithms for analyzing big data to get a better understanding of consumer behavior. The application of big data analysis has formed a new way of analyzing consumer behavior compared to the traditional way of thinking research by improving data resources compared to self-reported surveys (Hofacker et al., 2016). Since consumer behavior analysis heavily relies on data sources such as questionnaires, interviews, and databases, big data analytics enables companies to connect all essential elements to develop real-time approximations and conduct deeper analysis of consumer behavior (Dinu et al., 2016). For example, compared with interview or questionnaire field data, which is a little limited to data capture, attributes of big data in consumer behavior such as Google trends textual contents and photos can help researchers build a deeper understanding of consumers, such as their preferences and willingness to pay (Kanavos et al., 2018). In addition, as social networks rapidly expand, enabling

consumers to share their attitudes and comments, big data analytics and data mining technology have already significantly impacted consumer behavior analysis.

Along with the essential improvements big data and its relevant models and algorithms have brought to consumer behavior research, there are still some ongoing debates and challenges that need to be dealt with. The issue of class imbalance has garnered significant attention within the machine learning community in recent years. Despite the advancements in big data and deep learning, this problem persists (Rendon et al., 2020). Challenges still exist in selecting various data types and fully mining big data values (Bello-Orgaz et al., 2016). Moreover, as big data is often unstructured, there are still challenges to deeply analyzing and transforming data into meaningful information. Furthermore, data privacy issues impact internal and external stakeholders in various and potentially unexpected ways (Martin & Murphy, 2017).

A thorough review of the literature on this topic is necessary but currently unavailable, limiting our understanding of big data's role in consumer behavior research. Other studies have surveyed the literature on specific topics of big data research in consumer behavior, such as consumers repurchasing behavior (Shang & Li, 2017), behavior patterns (S. Singh & Yassine, 2019), product recommendation (Urkup et al., 2018), public attitude (A. Singh & Glińska-Newęs, 2022) and consumer decision-making behavior (Pai & Chen, 2023). Moreover, there is also some relevant literature covering certain aspects of big data research in consumer behavior, such as marketing research data (Green et al., 2020), user-generated content (Silva et al., 2020), web search data (Giglio et al., 2020), entreprise database (Adler et al., 2022), industry database (Volkova & Karpushkin, 2023), and professional database (Daviet et al., 2022).

However, two major gaps remain in this area. First, previous studies on big data focused heavily on models and algorithms such as data mining, data size, and visualization (Babiceanu & Seker, 2016; Ierkens et al., 2019), without providing enough insight into trends and key concerns in big data research. Second, as big data research rapidly evolves with new skills and techniques, past studies covered only a limited scope.

The following research questions are proposed by this study:

1. In the research field of consumer behavior, what types of big data, models, and algorithms are used?
2. What are the foremost relevant research topics and the new trends with big data and consumer behavior?
3. What are the influences, limitations, challenges, and future directions in the field of big data and consumer behavior?

To answer these research questions, we performed a comprehensive quantitative and qualitative systematic review of big data and relevant techniques in consumer behavior research, highlighting their current impact and the latest trends in this field. Our contribution aims to identify avenues for future research and advance the understanding of how big data can be leveraged to further improve consumer behavior research.

2. Data and Methods

Bibliometrics was introduced in the early 1990s and established formally in 1969 (A, 1969), then developed a crucial method for quantitative literature analysis (Diem & Wolter, 2013). With the analysis of authors, keywords, journals, countries, and institutions, bibliometrics helps trace the development of specific research fields (Abramo et al., 2011).

Advancements in computing have produced a better environment for bibliometric analysis through visualization. Research tools like CiteSpace and VOSviewer could generate knowledge graphs to get a deeper understanding of literature. CiteSpace utilizes set theory to standardize data and give visualization to reveal the evolution of research clusters, while VOSviewer applies probability theory for data standardization and gives many visualization options, for example, Network and Density to explore connections between authors and other elements (Eck & Waltman,

2009). The combination of the two tools could provide a framework for visualizing and excavating the developments and connections in a certain research area.

This paper follows the PRISMA protocol to make the flow chart and keep this paper replicability. The PRISMA protocol has been updated almost annually from 2009 to 2020. The PRISMA-2009 provides a 27-item checklist and a four-phase flow diagram to improve the systematic reviews and meta-analyses. The PRISMA statement gives guidance for reporting a systematic review (Moher et al., 2009). The PRISMA-2015 offers a 17-item checklist to facilitate the preparation of a robust protocol, which could raise the quality and reliability of systematic reviews (Moher et al., 2015). For PRISMA-2020, it has been introduced to update PRISMA-2009 to make a clear report of study selection, synthesis, and risk of bias by reflecting advancements in systematic review methodologies and technologies, for example, machine learning and natural language processing (Page et al., 2021). PRISMA-2009 aims to standardize and improve transparency in systematic reviews, and PRISMA-2020, with updates, takes advancements like machine learning and ensures systematic reviews remain reliable resources for researchers (Sarkis-Onofre et al., 2021).

To ensure the comprehensiveness and accuracy of the retrieval data, the summary of the data source is shown in Table 1, and Figure 1 represents the selection process followed by the PRISMA protocol (Page et al., 2021). The Web of Science index is selected as SCI-EXPANDED, SSCI, and the final Web of Science retrieval strategy is [TS = ((“big data” OR “data analytics” OR “data mining” OR “machine learning” OR “predictive analytics”) AND (“consumer behavior” OR “consumer behavior” OR “purchase behavior” OR “buying behavior” OR “shopping behavior” OR “customer behavior”))]. The Scopus search strategy is [TITLE-ABS-KEY ((“big data” OR “data analytics” OR “data mining” OR “machine learning” OR “predictive analytics”) AND (“consumer behavior” OR “consumer behavior” OR “purchase behavior” OR “buying behavior” OR “shopping behavior” OR “customer behavior”))]. The final search was (204+579=) 783 articles. The period is from January 2012 to December 2023, to cover the whole paper topics relevant to “big data” and “consumer behavior”. The document type is Articles, the language is selected as English, and the search date is January 14, 2024. Web of Science is exported as a full record, and Scopus is exported as all information. There are (133+238=) 371 articles after the primary limit time, article type, and language. After removing duplicate literature from 90 articles, there are still 281 articles. After filtering the titles and abstracts, 189 articles remained. The full text and article title abstract were further screened to obtain 127 valid articles.

Table 1. summary of data source.

| Category | Specific Standard Requirements |
|--------------------|--|
| Research database | Web of Science core collection, Scopus |
| Citation indexes | WOS (SSCI, SCIE), Scopus |
| Searching period | January 2012 to December 2023 |
| Language | “English” |
| Searching keywords | WOS |
| | TS = ((“big data” OR “data analytics” OR “data mining” OR “machine learning” OR “predictive analytics”) AND (“consumer behavior” OR “consumer behavior” OR “purchase behavior” OR “buying behavior” OR “shopping behavior” OR “customer behavior”)) |
| | SCOPUS |
| | TITLE-ABS-KEY ((“big data” OR “data analytics” OR “data mining” OR “machine learning” OR “predictive analytics”)) |
| | |

| | |
|--------------------|---|
| Subject categories | AND ("consumer behavior" OR "consumer behavior" OR "purchase behavior" OR "buying behavior" OR "shopping behavior" OR "customer behavior")) "Business" "Computer Science Information Systems" "Environmental Sciences" "Management" "Green Sustainable Science Technology" |
| Document types | "Articles" |
| Data extraction | Export with full records and cited references in RIS format |
| Sample size | 371 |

Source: Authors' elaboration.

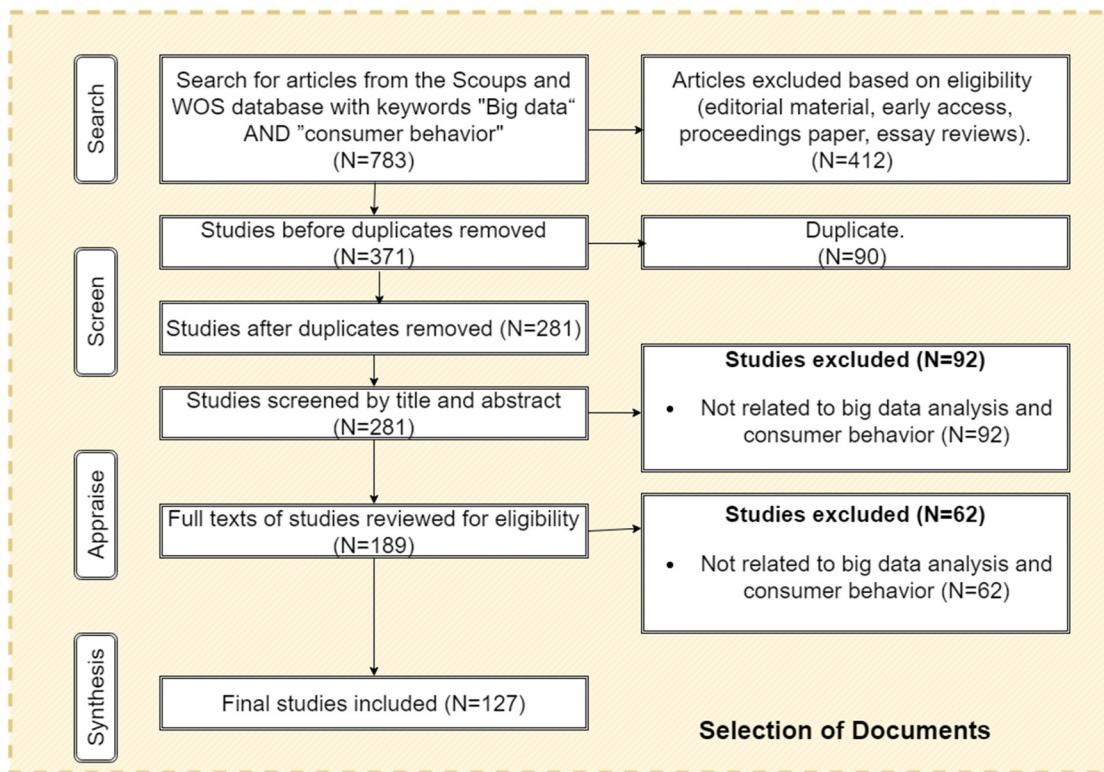


Figure 1. Process of literature selection. Source: Author's elaboration based on the selection process on Scopus and Web of Science.

After data cleansing, quantitative and qualitative analysis was performed on the final dataset of 127 research papers selected. This study employs a two-part methodology for the literature data analysis, starting with a bibliometric analysis to quantitatively assess the landscape of big data research in consumer behavior. In the second part, a thematic analysis is conducted, utilizing theme maps to qualitatively explore key themes and emerging trends within the literature.

From the studies that used big data to analyze consumer behavior, several major topics are associated with this topic. As shown in Figure 2, a theme map was built based on the literature analysis offers fundamental topics and clear relationships among big data literature in consumer behavior; this map can give a foundation for research in the future. Besides, the map could offer guidance for researchers to understand the state of big data research in consumer behavior recently and help provide future research directions.

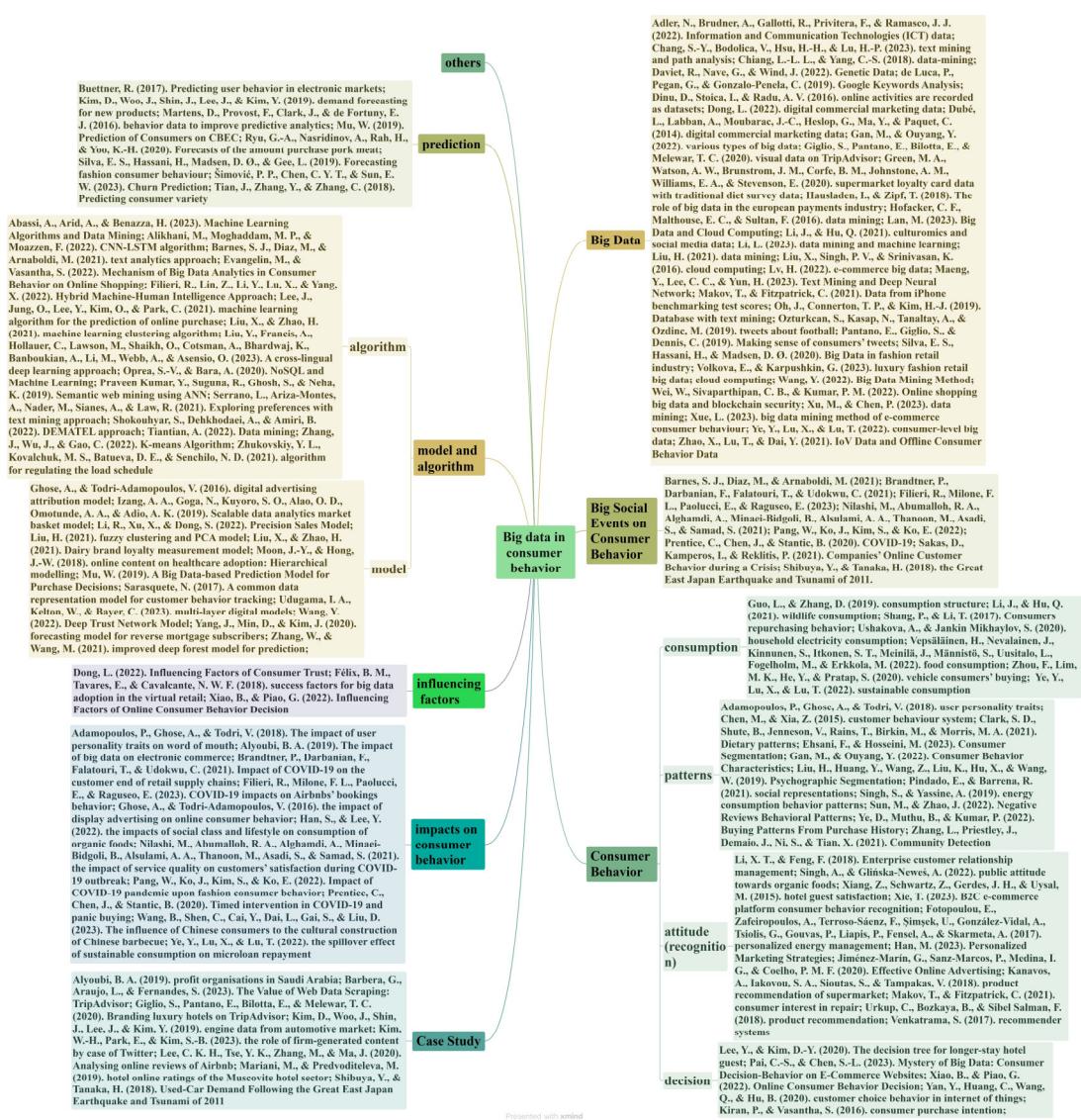


Figure 2. Literature theme map. Source: Author's elaboration of data collected from Scopus and Web of Science.

As illustrated in Figure 2, big data in consumer behavior analysis can be divided into several key categories. The primary category encompasses consumer behavior and big data types, with subcategories including consumption, patterns, preferences, attitudes, and decision-making processes. A detailed analysis of these big data and consumer behavior types is provided in the subsequent section of the study. Additionally, various models and algorithms, particularly those involving artificial intelligence (AI), are highlighted, with some AI algorithms demonstrating exceptional performance in recent years. Other topics, such as influencing factors and the impacts on consumer behavior, are also explored in the following sections.

3. Results of Bibliometric Analysis

The 127 papers used in this study were published in 110 journals by 83 authors from 284 organizations in 68 countries. Figure 3 presents the number of articles published every year on big data in the consumer behavior field. The first publication was in 2014. Starting from 2018, there was an increase in the number of publications. This trend could also be found in other studies (Chandra & Verma, 2023). The probable reason that big data in consumer behavior has increased rapidly after 2018 could be ascribed to the progress of technologies.

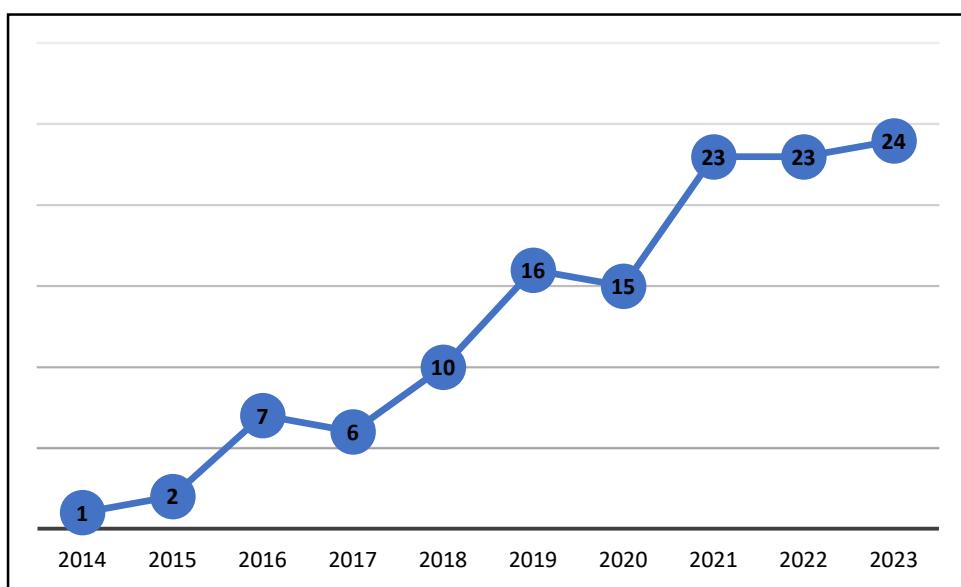


Figure 3. Distribution of articles by year of publication. Source: Author's elaboration of data collected from Scopus and Web of Science.

In terms of paper output by country, China, the United States, and South Korea published the most papers in this field, and the differences in paper output among other countries were relatively small. The top 4 countries (China, the United States, the Republic of Korea, and the United Kingdom) account for a significant portion of the total publications. A long tail distribution with many countries having fewer publications (2-3 each). European countries like Spain, Italy, Germany, and France have moderate numbers of publications. Other countries have fewer publications, indicating a potential for growth or differences in research focus and resources.

To better reflect the core authorship and relevance of big data in the field of consumer behavior analysis, 127 papers were visualized by author collaboration graph (Figure 5).

The node sizes in Figure 4 represent the number of papers published by the authors, and the lines represent the collaboration. There are 11 core authors in the sample literature, and the top authors are Lawson M Cade, Francis Azell, Webb Anne, Asensio Omar Isaac, Bhardwaj Khushi, Hollauer Catharina, Banboukian Aline, Cotsman Ashley, Shalkh Omar, and Li Mimi. In general, the cooperation in the field of Consumer big data field is relatively close, the research forces in this field are in a relatively concentrated state, and the scholars are closely connected, which is of great significance for the in-depth exploration of this field.

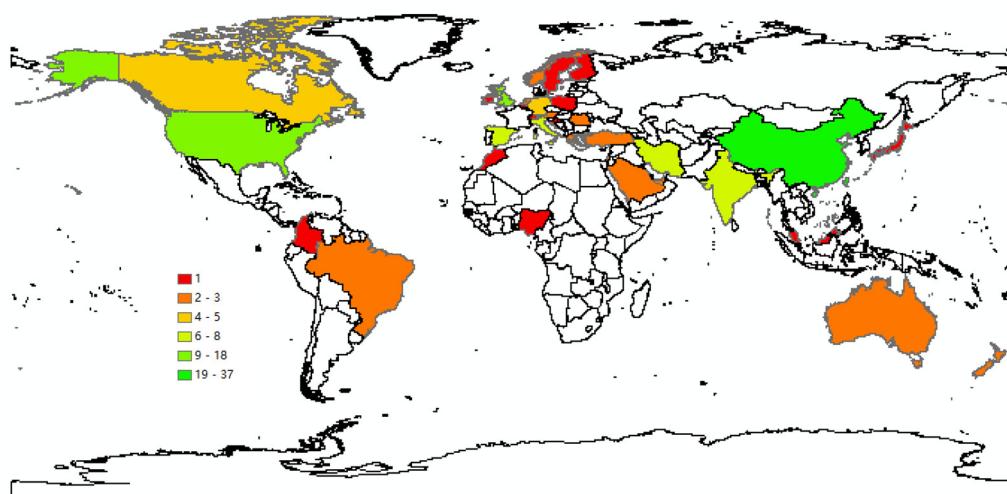


Figure 4. GIS map of high-yield countries.



Figure 5. Cooperation map of high-yield authors. Source of data: Author's elaboration of data collected from Scopus and Web of Science.

Using the “Detect Outbreak” function of Citespace software to observe deeper development changes, it was found that professional vocabulary has increased significantly in a short period in a specific year. Figure 6 is a prominent word map generated by the sample. From the highlighted text, the duration of hot spots shows a trend from long to short. From 2014 to 2018, the prominent keywords included marketing, behavior change, customer behavior, customer experience, and data analysis, with a duration of 7 to 9 years. From 2014 to 2017, the prominent keywords include genetic algorithms and growth, and the average duration of these keywords is as long as 4 years. Since 2018, academic research, machine learning, aging, consumer attitudes, and other issues have attracted

much attention. Among them, the keyword of machine learning has continued to this day, reflecting that related research will continue to be a key trend in the field of consumer big data in the future.

Top 10 Keywords with the Strongest Citation Bursts

| Keywords | Year | Strength | Begin | End | 2014 - 2023 |
|---------------------|------|----------|-------|------|-------------|
| marketing | 2014 | 1.56 | 2014 | 2018 | |
| behavior change | 2014 | 0.97 | 2014 | 2017 | |
| customer behaviour | 2015 | 1.38 | 2015 | 2018 | |
| big data analytics | 2016 | 1.85 | 2016 | 2019 | |
| customer experience | 2016 | 1.14 | 2016 | 2017 | |
| data analytics | 2017 | 1.95 | 2017 | 2018 | |
| academic research | 2018 | 0.97 | 2018 | 2019 | |
| machine learning | 2020 | 1.75 | 2020 | 2023 | |
| aged | 2020 | 1.3 | 2020 | 2021 | |
| consumer attitude | 2014 | 1.21 | 2020 | 2021 | |

Figure 6. Top 10 keywords with the strongest citation bursts. Source: Author's elaboration of data collected from Scopus and Web of Science.

Keywords reflect the interrelationships between the various topics represented in the literature and are the core summary of the article. The analysis of keywords is beneficial to the research of hot topics in this field. We run Citespace software to obtain the keyword co-occurrence map of consumer behavior research in the big data field, as shown in Figure 7. The size of each node in the graph represents the number of times the word appears. As the number of occurrences increases, the circle becomes larger. Among the keywords with a frequency of more than 20, big data, data mining, and behavior change appear the most. To sum up, there are nine main directions that scholars of consumer behavior research in the field of big data focus on: consumer, big data analysis, electronic commerce, commerce, consumer attitude, sales, big data, data mining and behavior change and other keywords, high Frequent keywords are the same as keywords that rank high in the center. That is, the higher the frequency, the more obvious the center. Centrality can cover hot spots and key turning points to a certain extent.

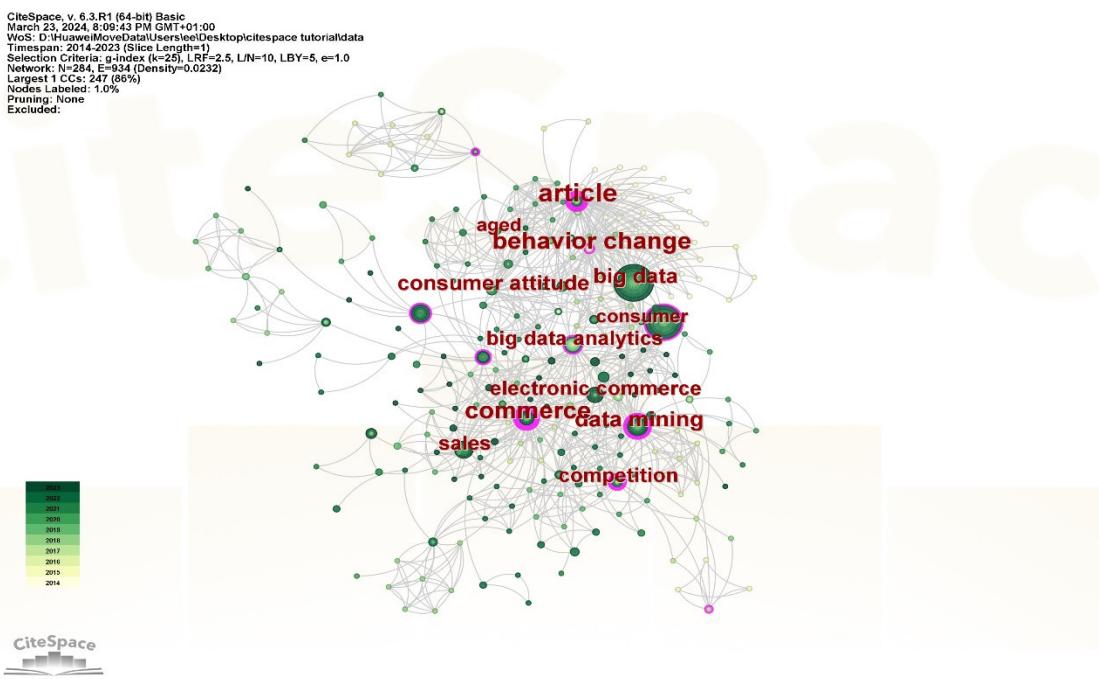


Figure 7. Keywords cluster map. Source: Author's elaboration of data collected from Scopus and Web of Science.

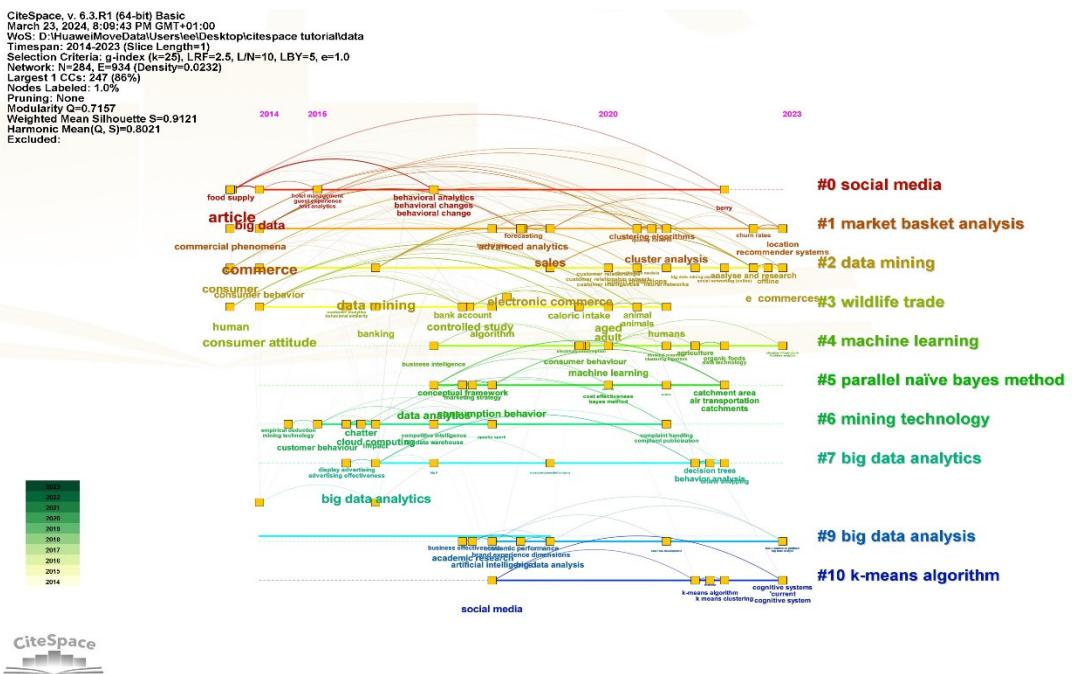


Figure 8. Keyword timeline chart. Source of data: Compiled by the author.

Combined with literature analysis, the keyword results are summarized into two categories. The first is algorithm research related to big data mining, which is the research on how computers acquire new knowledge and skills. On the one hand, it is directly used for information perception, data prediction, and estimation models. On the other hand, a new type of algorithm attempts to process and predict existing problems through a large amount of data learning accumulation. This model is also called "machine learning". Machine learning represents the forefront of human thinking and imitation technology. It can optimize certain systems by making predictions through training data sets, including model-based reinforcement learning, genetic-based multi-objective optimization

strategies, etc. Algorithms, data algorithms based on data mining, and feature analysis of support vector machines. The second is the change in consumer behavior and consumer attitudes, the change in consumer behavior for a certain category of specific agricultural products, or the impact of online consumer behavior. The third is interdisciplinary frontier disciplines, whose research fields are very broad. The innovative model of “big data + consumer behavior” will gradually mature in the development of technology and industry and achieve more new developments of cross-integration under the guidance of consumer market analysis, consumer preferences, consumer stickiness, and other needs.

4. Results of the Thematic Analysis

4.1. Types of Big Data Used in Consumer Behavior Research

Different types of big data have been used in consumer behavior research. As shown in Table 2, the data types can be separated into two categories by the content format, including structured and unstructured data. Structured data is highly organized and can be stored in a database like SQL, while unstructured data is complex to store. Then, these two categories can be split into six subcategories. The subcategories of unstructured data include (1) User-generated content, (2) web search data, and (3) marketing research data. The subcategories of structured data contain (1) Enterprise database, (2) Industry database, and (3) Professional database. In the following parts, the application of every big data type in consumer behavior research will be elaborated.

Table 2. Categories of big data used in consumer behavior research.

| Categories | Subcategories | Example articles |
|-------------------|---|--|
| Unstructured data | Marketing research data (e.g., Retailing and Advertising, diet survey data) | (Dubé et al., 2014), (Green et al., 2020) |
| | User-generated content (e.g., tweets, reviews of social websites, Google keywords) | (Ozturkcan et al., 2019), (Silva et al., 2020), (Pantano et al., 2019), (de Luca et al., 2019) |
| | Web search data (e.g., Taobao live banding data, TripAdvisor, e-commerce platform, Hadoop cloud computing platform) | (Xu & Chen, 2023), (Giglio et al., 2020), (Xue, 2023), (Wang & Zhang, 2021) |
| Structured data | Enterprise database (e.g., ICT dataset from airplane, EIS) | (Adler et al., 2022), (Upadhyay et al., 2024) |
| | Industry database (revenues of luxury) | (Volkova & Karpushkin, 2023) |
| | Professional database (e.g., genetic data, American Statistical Association DataFest) | (Daviet et al., 2022), (Y. Lee & Kim, 2020) |

Source of data: Compiled by the author.

Unstructured Big Data

In recent years, unstructured data has become a commonly used source of big data for analyzing consumer behavior. Researchers in consumer behavior analysis have utilized unstructured data collected from marketing research, user-generated content, and web search data to explore various issues related to consumer behavior. With the rapid development of the Internet and computer science, many types of unstructured big data have been generated.

For user-generated content data, the rise of the Internet has facilitated the growth of these social media platforms, allowing consumers to easily share their views online. These data sources provide valuable insights into consumer behavior; consequently, researchers have collected vast amounts of online reviews from social media and e-commerce platforms to identify consumer attitudes,

preferences, behavior patterns, and willingness to pay (e.g., Ozturkcan et al., 2019; Silva et al., 2020; Pantano et al., 2019; Xu & Chen, 2023). Understanding how brands can maintain the personalized and intimate relationship qualities provided by social media while meeting consumer expectations amid the increasing volume of interactions has become essential (Labrecque, 2014).

Other researchers use market research data to get a full understanding of availability, accessibility, affordability, and appeal in food consumption (Dubé et al., 2014). There's also a comparison between supermarket loyalty card data and traditional diet survey data for understanding the consumption behavior of protein in older adults in the UK (Green et al., 2020). Compared with User-generated data and Web search data, market research data allows researchers to focus on a certain type of questions in consumer behavior and make it more precise by questionnaires and interviews (Dabrowska, 2011). Much progress has been made in today's consumer behavior research using market research data. The new utilized netnography and MAXQDA software provided a new way for qualitative analysis, providing valuable insights for brands regarding consumer behavior on social media (Hosseini & Ghalamkari, 2018).

Finally, big data can be generated from consumer online browsing, searching, and buying behavior. Big Data has the potential to enhance the understanding of each stage in the consumer decision-making process. Traditionally, the field has advanced with a priori theory followed by experimentation. While the nature of the feedback loop between theory and results may shift significantly due to the influence of Big Data (Hofacker et al., 2016). Consumers usually search for information online, not only for comments from other consumers but also for product details and relevant information; they could even purchase online through e-commerce platforms. Compared to traditional data resources, web search data could provide a huge volume of numbers and text, which is useful for text mining and provides a comprehensive understanding of consumer attitudes, needs, and so on. For example, when it comes to examining the factors influencing consumers' online information search behavior with purchasing laptops and mobile phones, total search, number of searches, and cognitive ability were measured to understand consumer behavior (Dutta & Das, 2017). Besides, a personalized recommendation framework based on consumer web search data and the open-source Hadoop cloud computing platform was developed to enhance commodity exposure, recommend personalized products, and stimulate user consumption in e-commerce (Wang & Zhang, 2021). It could be concluded from these researchers jointly that the web search data provides more accuracy regarding consumer intention. Search queries (e.g., Xue, 2023; Xu & Chen, 2023) often directly reflect what consumers are interested in, planning to purchase, or researching at any given time. By analyzing these data, researchers can gain more precise insights into consumer behavior and trends, making them a valuable resource for predicting market dynamics and preferences.

Structured Big Data

A professional database, as identified in the literature, is limited in use but is a valuable resource maintained and updated by professional organizations. In consumer behavior research, these databases are advantageous because they provide structured data, eliminating the need for extensive data cleaning, which allows researchers to obtain and analyze data more conveniently. Advances in molecular genetics, for example, have led to rapid growth in the direct-to-consumer genetic testing industry, resulting in vast private genetic databases. Some researchers have examined the potential impact of this data on marketing by proposing a framework that integrates genetic influences into consumer behavior theory, exploring potential applications of genetic data in marketing (Daviet et al., 2022). While professional databases save time compared to other data sources, they have limitations in terms of flexibility and are typically restricted to specific research purposes. Similarly, industry databases also face limitations, particularly concerning topic specificity. For instance, a luxury industry database has been used in Russia to identify trends, patterns, and contradictions in demonstrative consumption within the fashion-retail sector, this includes addressing challenges related to digital marketing communications and e-commerce, as well as developing a systemic view

of big data within the marketing communications framework of Russia's fashion market (Volkova & Karpushkin, 2023).

Compared to professional and industry databases, enterprise databases are commonly used as structured data sources. For example, Google Trends, a public web tool by Google, Inc., based on Google Search, shows the frequency of specific search terms relative to the total search volume across various regions and languages. Both Google Trends and Wikipedia views data are structured numerically, unlike other data sources, which are typically in text format (X. Liu et al., 2016). In consumer behavior research, big data gathered from enterprise databases is often used to analyze various topics, such as consumer trends and market predictions. For instance, an ICT dataset from the aviation industry was used to create a Hotelling-inspired catchment area game, which analyzed the impact of collaboration between airports and airlines by integrating consumer behavior data with producers' financial data (Adler et al., 2022). Moreover, in the rapidly evolving landscape of digital marketing, leveraging data analytics within Enterprise Information Systems (EIS) has become crucial for businesses to better understand and engage with customers (Upadhyay et al., 2024).

4.2. Types of Consumer Behavior in Big Data Analysis

As shown in Table 3, previous studies have demonstrated that different types of big data have been utilized to analyze various aspects of consumer behavior, including consumption, patterns, preferences, attitudes, and decision-making. With the rise of social media platforms, e-commerce sites, and advanced algorithms such as AI, big data can now be collected more easily from these platforms and analyzed using the latest techniques (e.g., X. Liu et al., 2016; Adler et al., 2022). Unlike traditional studies that often relied on a single data source, big data from multiple origins can now be integrated and analyzed collectively to gain deeper insights into consumer behavior (Green et al., 2020). These findings indicate that data collected from collaborative datasets can provide a more comprehensive understanding of consumer behaviors when used supplementally.

The evolution of technology has given rise to various online social media platforms, such as Facebook, Twitter (now X), and TripAdvisor, providing consumers with avenues to share opinions and give feedback on products. As a result, consumers with similar buying patterns tend to gather and form groups (Adamopoulos et al., 2018). It is widely argued that the emergence and evolution of social media have significantly influenced how consumers obtain information and make decisions (Ghose & Todri-Adamopoulos, 2016). Since big data is generated from both pre-purchase and post-purchase activities, the information spread across websites offers an opportunity to gain a deeper understanding of consumption behaviors and consumer decision-making.

Table 3. Major themes of consumer behavior research in big data.

| Layers | Subject headings | Example articles |
|-------------|---|---|
| Consumption | consumption structure; wildlife consumption; Consumers repurchasing behavior; household electricity consumption; food consumption; vehicle consumers' buying; sustainable consumption | Guo, L., & Zhang, D. (2019); Li, J., & Hu, Q. (2021); Shang, P., & Li, T. (2017); Ushakova, A., & Jankin Mikhaylov, S. (2020); Vepsäläinen, H., Nevalainen, J., Kinnunen, S., Itkonen, S. T., Meinilä, J., Männistö, S., Uusitalo, L., Fogelholm, M., & Erkkola, M. (2022); Zhou, F., Lim, M. K., He, Y., & Pratap, S. (2020); Ye, Y., Lu, X., & Lu, T. (2022). |

| | | |
|------------------------|--|--|
| Patterns | user personality traits; customer behaviour system; Dietary patterns; Consumer Segmentation; Consumer Behavior Characteristics; Psychographic Segmentation; social representations; energy consumption behavior patterns; Negative Reviews Behavioral Patterns; Buying Patterns from Purchase History; Community Detection | Adamopoulos, P., Ghose, A., & Todri, V. (2018); Chen, M., & Xia, Z. (2015); Clark, S. D., Shute, B., Jenneson, V., Rains, T., Birkin, M., & Morris, M. A. (2021); Ehsani, F., & Hosseini, M. (2023); Gan, M., & Ouyang, Y. (2022); Liu, H., Huang, Y., Wang, Z., Liu, K., Hu, X., & Wang, W. (2019); Pindado, E., & Barrena, R. (2021); Singh, S., & Yassine, A. (2019); Sun, M., & Zhao, J. (2022); Ye, D., Muthu, B., & Kumar, P. (2022); Zhang, L., Priestley, J., Demaio, J., Ni, S., & Tian, X. (2021). |
| Attitude (recognition) | Enterprise customer relationship management; public attitude towards organic foods; hotel guest satisfaction; consumer behavior recognition; personalized energy management; Personalized Marketing Strategies; Effective Online Advertising; consumer interest in repair; product recommendation; recommender systems | Li, X. T., & Feng, F. (2018); Singh, A., & Glińska-Newes, A. (2022); Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015); Xie, T. (2023); Fotopoulou, E., Zafeiropoulos, A., Terroso-Sáenz, F., Şimşek, U., González-Vidal, A., Tsiolis, G., Gouvas, P., Liapis, P., Fensel, A., & Skarmeta, A. (2017); Han, M. (2023); Jiménez-Marín, G., Sanz-Marcos, P., Medina, I. G., & Coelho, P. M. F. (2020); Kanavos, A., Iakovou, S. A., Sioutas, S., & Tampakas, V. (2018); Makov, T., & Fitzpatrick, C. (2021); Urkup, C., Bozkaya, B., & Sibel Salman, F. (2018); Venkatrama, S. (2017). |
| Decision | decision tree; Online Consumer Behavior Decision; customer choice behavior in internet of things; consumer purchase intention | Lee, Y., & Kim, D.-Y. (2020); Pai, C.-S., & Chen, S.-L. (2023); Xiao, B., & Piao, G. (2022); Yan, Y., Huang, C., Wang, Q., & Hu, B. (2020); Kiran, P., & Vasantha, S. (2016). |

Source of data: Compiled by the author.

The following paragraphs will carefully analyze the detailed subcategories of consumer consumption, attitudes, patterns, and predictions of consumer behavior.

Consumer Consumption

As shown in Table 3, the literature review covers various topics, including reasons for consumer consumption (F. Zhou et al., 2020), sustainable consumption (Y. Ye et al., 2022), consumption structure (Guo & Zhang, 2019), food consumption (Vepsäläinen et al., 2022), repurchase (Shang & Li, 2017), and household electricity consumption (Ushakova & Jankin Mikhaylov, 2020).

For text mining analysis, the process involves content filtering and data extraction (J. Li & Hu, 2021). In one study, consumers were categorized into green consumers and traditional consumers, the findings highlighted the potential of sustainable consumption to generate positive spillover effects across various domains of consumer behavior. (Y. Ye et al., 2022).

For data mining analysis, researchers found that “cost-effectiveness” is the most critical factor influencing vehicle consumers, indicating that consumers prefer to purchase vehicles with favorable pricing (F. Zhou et al., 2020). Analysis of e-commerce platform data suggests that consumption

structures vary across different periods, showing a trend toward consumption upgrading (Guo & Zhang, 2019). Grocery purchase data are considered moderately valid for describing food consumption patterns among adult populations (Vepsäläinen et al., 2022). Additionally, residential smart meters can effectively reveal energy consumption patterns (Ushakova & Jankin Mikhaylov, 2020).

Moreover, leveraging social media and search engine data helps in developing more effective strategies for wildlife conservation by addressing wildlife consumption (Li & Hu, 2021). Traditional database technologies struggle to process large volumes of static customer information and dynamic transaction data efficiently, whereas machine learning techniques offer effective solutions to this challenge (Li & Hu, 2021).

Consumer Attitude

The emergence of data analysis technologies and algorithms, particularly artificial intelligence (AI), has created a need for analyzing consumer attitudes and recognition. Some researchers utilize big data and text mining to analyze hotel guest satisfaction (Xiang et al., 2015) and public attitudes toward organic food (A. Singh & Glińska-Newęs, 2022). For consumer recognition, researchers employ AI to identify consumers, and enterprises use data mining to manage relationships with them (Xie, 2023; X. T. Li & Feng, 2018). In summary, big data mining and AI are commonly used methods in the processes of consumer recognition and the identification of consumer attitudes. The reviewed literature indicates that user-generated content (UGC) is predominantly used as a type of big data to investigate consumer attitudes, particularly focusing on consumer preferences and satisfaction. Unlike traditional survey-based research that relies on "stated preferences" to assess satisfaction levels (Lv et al., 2022), big data research commonly utilizes UGC (e.g., online reviews, tweets, e-commerce platform comments) as it provides larger samples, yields better results, and reduces potential subjective biases. Previous studies have also identified price, subsidies, and after-sales service as key factors influencing consumer preferences, with price playing a particularly dominant role (Jung et al., 2021). Additionally, researchers found that mental depreciation and perceived value significantly affect consumer attitudes toward smartphone repairs (Makov & Fitzpatrick, 2021).

In addition, when comparing traditional analysis with big data analysis of consumer preferences, traditional methods often treat consumer preference data subjectively through self-rated questionnaires or interviews. In contrast, big data analysis tools primarily evaluate data attributes objectively using predefined rules (Makov & Fitzpatrick, 2021). For example, with the use of big data and AI technologies, enterprises can gain a deeper understanding of consumer preferences, leading to personalized marketing strategies. Internet of Things (IoT) technologies also facilitate data collection through various devices, enabling a more comprehensive understanding of consumer preferences (M. Han, 2023).

Regarding consumer preferences, researchers have primarily focused on product recommendations and personalized management. In the field of personalized management, particularly in the energy sector, identifying key energy-consuming factors is crucial. An energy-aware IT ecosystem has been established to collect relevant information and develop personalized management using modern techniques such as IoT, data modeling, and personalized recommendation mechanisms (Fotopoulou et al., 2017). An empirical analysis concluded that spatial-temporal mobility and financial features are significant factors in predicting consumer attitudes (Urkup et al., 2018). Through IoT, enterprises can obtain diverse data and gain a deeper understanding of consumer preferences, which is essential for personalized marketing strategies (M. Han, 2023). For product recommendations, a large-scale supermarket product recommendation system was developed based on consumer behavior using data mining techniques from Amazon (Kanavos et al., 2018). This analysis demonstrated the effectiveness of cloud infrastructure and MapReduce as a programming environment. Moreover, research on consumer preferences can enhance advertising effectiveness; companies are advised to estimate the number of customers and

potential sales opportunities based on consumer preferences (Jiménez-Marín et al., 2020). An intelligent recommendation framework was also developed for consumer recommendation systems and has been applied in B2C e-commerce scenarios (Fotopoulou et al., 2017).

Consumer Patterns

Identifying and clarifying consumer patterns has been a significant concern for predicting consumer purchasing tendencies (Adamopoulos et al., 2018). Researchers have utilized various data sources, such as tweets (Pindado & Barrena, 2021) and supermarket transaction data (Clark et al., 2021). Analytical techniques, including cloud computing (D. Ye et al., 2022), text mining (Adamopoulos et al., 2018), and data mining (M. Chen & Xia, 2015), are commonly employed. Big data analysis processes can be categorized into two primary methods: text analysis and data analysis.

In text analysis, research on tweets has found that consumer patterns vary across regions, with regional cultural contexts significantly influencing users' attitudes toward food innovations (Pindado & Barrena, 2021). Text mining of word-of-mouth (WOM) messages revealed a positive and statistically significant effect of personality similarity between social media users on the likelihood of subsequent purchases after exposure to WOM messages (Adamopoulos et al., 2018). Additionally, researchers identified psychographic segmentation using natural language processing (NLP) methods based on two online psychographic lexicons: the Big Five Factor (BFF) personality traits and the Schwartz Value Survey (SVS), both derived from users' word usage (H. Liu et al., 2019). Consumer density is utilized to estimate geographical peculiarities (Clark et al., 2021). Analysis of online negative reviews has also provided insights into online consumer temporal, perceptual, and emotional patterns (Sun & Zhao, 2022).

In data analysis, researchers have used various techniques to identify tourists' characteristics, revealing differences in secondary consumption behaviors among different types of tourists (Gan & Ouyang, 2022). Additionally, an unsupervised progressive incremental data mining mechanism was proposed to extract and analyze energy consumption patterns using frequent pattern mining methods (S. Singh & Yassine, 2019). By accurately identifying customer segments with similar purchasing preferences, actionable cross-selling strategies were enhanced, ultimately increasing consumer loyalty (L. Zhang et al., 2021). Supermarket sales data serve as a valuable resource for large-scale dietary research and can be utilized to clarify public dietary patterns (Clark et al., 2021).

Consumer Decision

Regarding consumer decision-making, one key topic is consumer willingness to buy, which is analyzed through user-generated content (UGC). UGC is crucial for understanding consumer intentions. Research has shown that UGC generated from social media and big data significantly impacts consumer purchase intentions (Kiran & Vasantha, 2016). Another related topic is online shopping decisions. Researchers have found that reducing touchpoints on e-commerce platforms can enhance revenue conversion rates (Pai & Chen, 2023), and the internet helps consumers better understand the products or services offered by companies (Xiao & Piao, 2022). The Internet of Things (IoT) has transformed traditional consumer networks. To address consumer choice mechanisms and reduce confusion, IoT and big data analysis can be highly effective (Yan et al., 2020).

Moreover, a decision tree model and artificial intelligence (AI) were used to analyze consumer behavior on Expedia.com. The analysis found that longer-stay hotel guests tend to prioritize package deals, and their preferences vary by region (Y. Lee & Kim, 2020).

Predictions of Consumer Behavior

After identifying consumer patterns and attitudes, predicting future consumer behavior is essential for companies to develop more effective marketing strategies. Accurate forecasting not only enhances consumer management but also contributes to the economic growth of companies (Buettner, 2017). Commonly used data for predicting consumer behavior include web search data,

such as the XING dataset from online social networks (Buettner, 2017), also data related to consumer behavior and payments (Martens et al., 2016), Google Trends (Silva et al., 2019), customer relationship management datasets (Šimović et al., 2023), and cross-border e-commerce platforms (Mu, 2019). Additionally, data analysis often involves a combination of structured and unstructured data (Ryu et al., 2020).

The collected 127 consumer behavior studies on forecasting using big data can be traced back to 2016 (Martens et al., 2016). Most of these studies have concentrated on forecasting personality-based product preferences (Buettner, 2017), consumer variety-seeking behavior (Tian et al., 2018), and fashion consumer trends (Silva et al., 2019), among others. By utilizing both structured and unstructured data from diverse sources (e.g., Google Trends, tweets, dietary survey data), AI algorithms have significantly improved the accuracy of forecasting consumer behavior compared to traditional methods such as linear regression and regression trees (Jackson & Ivanov, 2023). Furthermore, web search data is widely used by researchers to predict consumer trends; specifically, data from Google Trends has been shown to provide more reliable forecasts of tourist behaviors (Havranek & Zeynalov, 2021). These studies have not only enhanced the prediction of consumer trends, which is valuable for marketing, but also validated algorithms relevant to machine learning and text mining.

4.3. Application of Models and Algorithms in Big Data Usage

The advent of digital transformation, social network adoption, cloud infrastructure, and big data technology now enables researchers to develop models to track and store observed customer behavior (Sarasquete, 2017). A Big Data-based Purchase Decisions Prediction Model was established to analyze consumer behavior on cross-border e-commerce platforms (W. Mu, 2019). An Improved Deep Forest model was designed to predict e-commerce consumers' repurchase behavior (W. Zhang & Wang, 2021). Beyond the prediction of purchase and repurchase behaviors, a Deep Trust Network Model was developed to understand behavioral characteristics in consumer patterns (Y. Wang, 2022). For predicting consumer behavior, a Neural Network-based Precision Marketing Model focuses on data mining to study user churn prediction and user value enhancement, which are the two most critical factors influencing marketing revenue (H. Liu, 2021). In the context of product lifecycle management, a multi-scale digital model was developed to support complex decision-making (Udugama et al., 2023).

As big data technology rapidly evolves, new algorithms and models have emerged to handle the extensive use of large datasets. For example, text mining and data mining play significant roles in this field (Praveen Kumar et al., 2019). These algorithms are instrumental in predicting and tracking consumer behavior. An algorithm for collecting and processing power consumption data, along with a load planning algorithm, was developed to encompass all levels of device interaction (Zhukovskiy et al., 2021). Furthermore, consumer preferences and attitudes (Serrano et al., 2021), online consumption behavior (Evangelin & Vasantha, 2022), and energy consumption (Abassi et al., 2023) were all analyzed using text mining and data mining, making these algorithms widely popular in the analysis of big data for consumer behavior. In addition to data mining and text mining, CNN-LSTM (Alikhani et al., 2022) and ANN (Praveen Kumar et al., 2019) models have been applied in demand response and retail management.

Consumer behavior researchers have also employed AI algorithms in big data analysis for consumer behavior research. Among AI algorithms, machine learning and deep learning are commonly used in this context. An analysis of online consumer behavior data from the Google Merchandise Store found that the ensemble model, eXtreme Gradient Boosting (an AI algorithm), is the most suitable for predicting purchase conversions among online consumers, with oversampling identified as the best method to mitigate data imbalance bias (J. Lee et al., 2021). Additionally, researchers have demonstrated that machine learning is more effective than traditional algorithms, with larger sample sizes leading to more accurate results. The decision tree model and the arithmetic

mean calculation method have proven more effective than conventional algorithms (R. Li et al., 2022). The combination of machine learning and data mining has also been utilized to analyze consumer behavior in the context of energy usage (Abassi et al., 2023). Thus, it is likely that new AI algorithms and techniques will continue to be adopted in consumer behavior research in the future.

4.4. Other Research Themes of Big Data in Consumer Behavior

As shown in Figure 9, other research topics could be characterized into case study, influencing factors, impacts on consumer behavior, and big social events on consumer behavior.

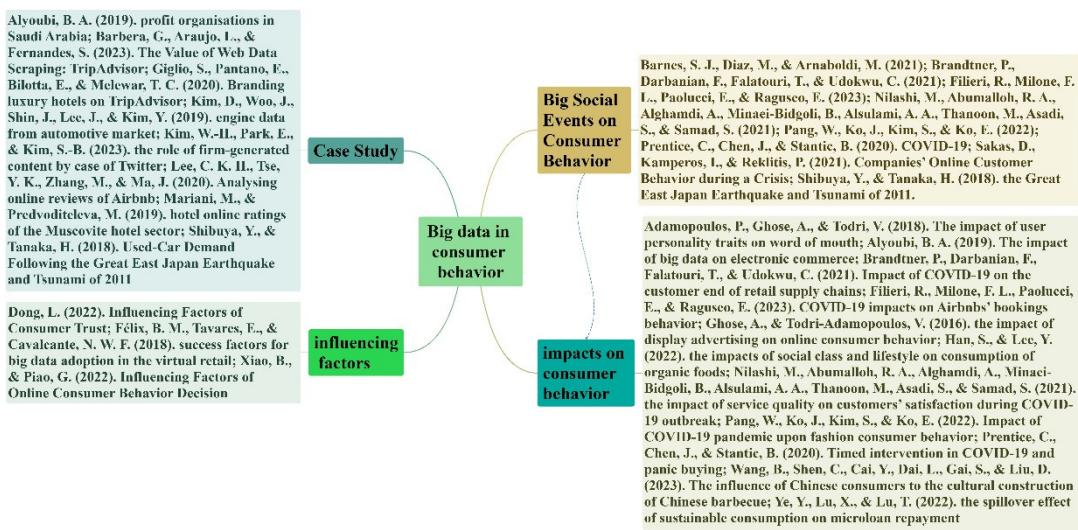


Figure 9. Other themes relevant to big data and consumer behavior. Source of data: Compiled by the author.

Influencing Factors

In the 127 papers reviewed, only three were directly related to the influencing factors of consumer behavior utilizing big data, focusing on online shopping characteristic data (Xiao & Piao, 2022), e-commerce marketing platform data for green agricultural products (L. Dong, 2022). Some of these studies also combined other sources of big data. It was concluded that consumer decision-making behavior occurs in stages, with different factors influencing each stage of the process (Xiao & Piao, 2022). Product factors, psychological factors, income factors, social factors, and cultural factors are the main determinants of consumer behavior (L. Dong, 2022). Additionally, privacy management, strategic alignment, structure, and functions influence the adoption of big data (Félix et al., 2018).

Impacts on Consumer Behavior

Consumer behavior can be impacted by a wide range of factors. For instance, cultural influences can significantly affect consumer behavior. A study in Korea found that while social class characteristics influence the consumption of organic foods, individual lifestyle plays a more critical role in actual purchasing behavior (S. Han & Lee, 2022). Cultural factors also impact the consumption of Chinese barbecue, as regional similarities in consumer preferences for barbecue continue to grow (B. Wang et al., 2023). Besides culture, researchers have found that sustainable consumption is more prevalent among consumers who prefer socially oriented products (Y. Ye et al., 2022). Personality traits also impact consumer behavior; consumers who exhibit similarity on social media platforms are more likely to make purchases (Adamopoulos et al., 2018). Additionally, consumers who are less exposed to display advertising have an increased tendency to search for brands, while those with higher exposure to display advertising are more likely to engage in direct search behavior (Ghose & Todri-Adamopoulos, 2016).

Big Social Events on Consumer Behavior

Among the 127 selected papers, eight focused on the impact of major social events. As shown in Figure 8, there is a clear relationship between major social events and their impact on consumer behavior. For example, many researchers have examined the impacts of COVID-19 on various aspects of consumer behavior, including Airbnb booking behavior (Filieri et al., 2023), companies' online consumer behavior (Sakas et al., 2021), consumer satisfaction at the point-of-sale (Brandtner et al., 2021), and differences between mass and luxury products (Pang et al., 2022). The pandemic also influenced panic buying (Prentice et al., 2020; Barnes et al., 2021) and service quality perception (Nilashi et al., 2021). Additionally, after the Great East Japan Earthquake and Tsunami, social media related to rebuilding activities was found to be positively correlated with the demand for used cars (Shibuya & Tanaka, 2018).

Case Study

There are eight case studies among the 127 papers reviewed. One case study on the 2011 Great East Japan Earthquake and Tsunami was conducted to understand car scarcity and demand in the aftermath, finding that social media interactions related to rebuilding efforts and emotional support were positively correlated with increased demand for used cars (Shibuya & Tanaka, 2018). Another case study analyzed online reviews on platforms like Airbnb (C. K. H. Lee et al., 2020), branding luxury hotels on TripAdvisor (Giglio et al., 2020; Barbera et al., 2023), and Muscovite hotels (Mariani & Predvoditeleva, 2019), along with firm-generated content from Twitter (now X) (W.-H. Kim et al., 2023). Furthermore, researchers used data obtained from the automotive market to improve the accuracy of new product demand forecasting (D. Kim et al., 2019).

5. Conclusion

With the rapid growth of new information and technologies in the field of big data, various types of large-scale data have been tested, analyzed, and filtered to generate new insights into consumer behavior. This study offers a systematic review of the application of big data in consumer behavior research, highlighting key research topics. Given the distinct characteristics of structured and unstructured data, which are used independently or in combination to address problems, a systematic literature review is essential for gaining a comprehensive understanding of recent studies and offering insights for future research.

5.1. Main Findings and Implications

This research identified 127 journal articles published between 2012 and 2023. As shown in Table 1, the bibliometric analysis in Section 3 indicates that leading journals in consumer behavior have increasingly published studies incorporating big data analysis contributed by researchers worldwide. Although studies on big data in consumer behavior have gained attention, the field remains in its early stages, as the first identified paper in this review dates to 2014. Since 2014, big data analysis in consumer behavior has shown an upward trend in publication frequency, increasing from 1 paper in 2014 to 24 papers in 2023. The systematic review of these articles addressed the gaps and questions raised in the introduction.

First, big data in consumer behavior research is categorized into structured and unstructured data. Structured data includes enterprise databases, industry databases, and professional databases, while unstructured data encompasses traditional marketing research data, user-generated content, and web search data. Some researchers argue that combining structured and unstructured data is more effective. Second, in the application of models and algorithms, it is noted that AI in big data analytics can significantly influence traditional consumer behavior analysis. Some researchers have tested various AI models and algorithms to identify the most suitable approaches for predicting consumer behavior, particularly in forecasting consumption trends. The primary AI methods

identified are deep learning and machine learning. Besides AI algorithms, the primary big data analysis methods are text mining and data mining. Third, this review identified major types of consumer behavior analyzed through big data, including consumer consumption, attitudes, patterns, and decisions, with a focus on behavior predictions. These aspects of consumer behavior have been extensively studied using big data analysis, yielding significant insights and contributions.

These findings have valuable implications for academics. On the one hand, with the rapid advancement of modern technologies, big data has garnered significant attention in consumer behavior research. However, few studies have summarized and updated the trends in this field. This systematic review serves as a guideline for consumer behavior researchers to understand current trends in big data analytics and provides insights, along with identifying research gaps for future studies. On the other hand, this review highlights the role of big data in consumer behavior research and clarifies the main types of big data and consumer behaviors studied in recent years. This contributes to the field by enhancing the understanding of big data's impact on consumer behavior analysis.

The findings also have practical implications for companies and enterprises. For example, by leveraging big data and advanced AI algorithms to predict market trends and consumer preferences, companies can better anticipate trends and develop future strategies. For instance, research has shown that social data can be used to predict a user's personality (Buettner, 2017). Additionally, by analyzing consumer patterns, companies can create personalized recommendations and advertisements, enhancing their development. Furthermore, by assessing consumer attitudes, companies can gauge product acceptance, leading to improved product management. It is suggested that hotel guest experiences should be analyzed more granularly and nuancedly, enabling hotel managers to use these insights to enhance guest satisfaction by focusing on key dimensions of the guest experience (Xiang et al., 2015). For managers, this review may provide new perspectives on managing their business processes. Undoubtedly, the advent of big data, along with relevant models, algorithms, and AI, provides a new angle on understanding consumer behavior based on consumer-generated content. However, extracting valuable information from this data requires relevant knowledge and a fresh perspective. This review offers potential guidance for managers to pinpoint key aspects of big data in consumer behavior and identify its potential value. Moreover, this review can also help consumers understand the rapid development of big data in consumer behavior and encourage them to benefit from services enhanced by big data and AI technologies.

5.2. Challenges and Future Directions of Big Data Research in Consumer Behavior

Challenges

Although there has been progress in big data analysis, challenges remain regarding data handling, suitable data analysis methods, data reliability, data privacy, and legal issues.

First, since big data requires large volumes of datasets or data sources, advanced technology and high costs are often involved. Unstructured data, particularly user-generated content, necessitates efficient methods for identification, collection, and filtering. Some studies require a combination of structured and unstructured data (Ryu et al., 2020).

Second, the primary challenge today is finding the most suitable models and algorithms for analyzing specific questions. Some researchers have developed models to address specific questions and demonstrated their effectiveness, while others are still testing the efficiency of current AI models and algorithms. It has been demonstrated that the ensemble model, eXtreme Gradient Boosting (XGBoost), is particularly effective for predicting online consumer purchase conversion, with oversampling being the most effective method to address data imbalance (J. Lee et al., 2021).

Third, the quality of collected data remains questionable. Some researchers have developed frameworks to detect fake reviews in online consumer electronics retailers (Barbado et al., 2019). Additionally, some user-generated content is too subjective and may not accurately reflect the true quality of products. Furthermore, although big data analytics has significantly advanced hospitality

and tourism research, there are growing concerns about data quality, particularly with user-generated content (UGC). For instance, consumers might post fake or low-quality reviews online, and some consumer-generated data may be subjective and fail to reflect objective realities (Xue, 2023).

Moreover, some organizations may generate fake big data to attract customers (e.g., spammers hired to post favorable comments on official social media pages). Therefore, the reliability and validity of big data present significant challenges for researchers.

Finally, privacy concerns are a common and significant challenge in the era of big data. For individuals, some types of big data involve sensitive personal information, such as phone numbers, bank account details, and transaction information. For organizations, certain types of big data may also include trade secrets, such as internal databases containing customer information. Although this information is crucial for studying hospitality and tourism management, obtaining it is challenging due to privacy concerns. This difficulty has become a notable obstacle in hospitality and tourism research. For instance, researchers reported that they were unable to explore how visitors' behaviors differ across various party sizes due to restricted access to mobile tracking data containing personal information (Zhao et al., 2021).

To address these challenges related to data capture and analysis, more advanced AI techniques are required to obtain more detailed and accurate insights from representative samples. Additionally, enhancing collaboration between academia and industry may offer a practical approach to addressing the challenges of data reliability and privacy.

Future Directions

Although research on big data and AI has already made significant contributions to understanding consumer behavior, there remains substantial room for future development—particularly in the expansion of data types, the refinement of analytical techniques, and the exploration of a broader set of research issues. In particular, artificial intelligence (AI) and machine learning (ML) should play a more central role in addressing these research gaps.

Most current studies, especially in the fields of hospitality and tourism, rely heavily on single-source datasets such as online reviews or official government statistics. While these sources are valuable, dependence on a single type of data may limit the depth and reliability of insights. Recent findings indicate that combining multiple data sources including structured and unstructured data yields more accurate, objective, and comprehensive results (Anderson et al., 2016, Chaudhary et al., 2021). Therefore, future research should prioritize integrating heterogeneous datasets and apply AI-driven data fusion techniques to enable more nuanced consumer behavior analysis. In addition, combining big data with traditional methods such as surveys, in-depth interviews, and experiments can further strengthen the robustness and validity of findings.

Moreover, analytical techniques must evolve to meet the demands of increasingly complex and diverse data. AI and ML methods, especially deep learning, natural language processing (NLP), and ensemble models like XGBoost, offer powerful capabilities for extracting hidden patterns, predicting future behavior, and managing large-scale unstructured data. These technologies are particularly effective in enhancing personalization, forecasting trends, and identifying sentiment in consumer-generated content. Future research should not only refine the application of these algorithms but also explore hybrid approaches that combine multiple AI techniques to improve interpretability and accuracy (C. L. P. Chen & Zhang, 2014; Y. Wang et al., 2018).

Future studies should also expand the scope of inquiry to cover more contemporary and diverse issues. At the individual level, residents and hospitality and tourism employees are important subjects for investigation. In the big data era, their lives and work have been significantly impacted by various events, providing valuable information to better understand their interactions with guests. For example, researchers focused on public perceptions of robots as hotel frontline employees by analyzing online reviews (Xiang et al., 2015). Additionally, interactions among different stakeholders should be explored in future research (Fang et al., 2020). Few studies have focused on the group level

in the selected articles. However, researchers noted that different travel group compositions (e.g., family travelers, couples) exhibit varying attitudes and behaviors (Vepsäläinen et al., 2022). Therefore, future research should focus more on the group level to explore attitudes (e.g., group experience, group satisfaction) and behaviors (e.g., group booking, group decision-making) of various travel groups. Some researchers also advocated for increased research at the group level (Ierkens et al., 2019). At the organizational level, while previous research has primarily focused on for-profit organizations, nonprofit organizations warrant more attention. Nonprofit organizations possess vast amounts of big data and face challenges in managing, analyzing, and applying it to enhance service and consumer experience (Grandhi et al., 2021). Therefore, it is worthwhile to investigate nonprofit organizations using innovative big data and analytics. Furthermore, other organizational functions, such as human resource management (HRM), require more exploration beyond marketing and performance analysis.

Finally, at the industry level, other significant issues (e.g., events, transportation) should also be explored using systematic big data analysis. For instance, the management of major events (e.g., festivals, disasters) that can have a substantial impact on industry development warrants further investigation. For example, when a natural disaster occurs in a scenic area (e.g., the earthquake in Nine-Village Valley, Sichuan, China), it often leads to extensive discussions and information sharing on social media platforms. Thus, the rich data generated from these platforms can be used to identify critical issues and track public opinion trends. Furthermore, big data has not only disrupted the development of the hospitality and tourism industry but has also impacted other sectors. Therefore, future research should also examine cross-industry interactions and integrations. Additionally, it is important to note that existing research primarily focuses on the positive aspects of big data in hospitality and tourism research and practice. The negative aspects of big data usage remain largely unexplored and should be investigated in future research.

5.3. Limitations and Future Directions of This Study

Although this study presents a systematic and rigorous review of big data analysis and the application of AI in consumer behavior research, several limitations should be acknowledged. First, the review exclusively focused on peer-reviewed journal articles, excluding other potentially valuable sources such as conference proceedings, books, reports, and review articles. Future reviews could benefit from incorporating a broader range of literature to capture a more holistic view of the field. Second, this study only included publications written in English, which may limit the diversity of perspectives, particularly given the global nature of consumer markets. Future research should consider including non-English sources to enhance cultural and contextual comprehensiveness. Third, the selection process was conducted manually, which may have introduced selection bias despite efforts to ensure objectivity and consistency. Employing AI-based tools and algorithms for literature screening and classification in future reviews could improve efficiency, reduce bias, and enhance the reproducibility of findings. Finally, as the big data landscape continues to evolve rapidly, future systematic reviews should consider dynamic updating mechanisms—potentially through automated literature mining tools—to keep pace with emerging trends and technologies in consumer behavior research.

Author Contributions: **Qiankun Liu:** Data collection and Visualization, Original draft preparation, and revisions. **Ruigang Wang:** Paper revisions. **Muhabaiti Pareti:** Paper guidance and review. **Alessandra Castellini:** Paper guidance and review. **Maurizio Canavari:** Paper guidance and review.

Funding: This work was supported by the Fundamental Research Funds for the University of Bologna and the Chinese Scholarship Council (CSC) No.202308650019.

Acknowledgment: This work was supported by the Fundamental Research Funds for the University of Bologna and the Chinese Scholarship Council (CSC) No.202308650019.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

A, P. (1969). Statistical bibliography or bibliometrics. *Journal of Documentation*, 25, 348.

Abassi, A., Arid, A., & Benazza, H. (2023). Moroccan Consumer Energy Consumption Itemsets and Inter-Appliance Associations Using Machine Learning Algorithms and Data Mining Techniques. *Journal of Engineering for Sustainable Buildings and Cities*, 4(1). Scopus. <https://doi.org/10.1115/1.4062113>

Abramo, G., D'Angelo, C. A., & Viel, F. (2011). The field-standardized average impact of national research systems compared to world average: The case of Italy. *Scientometrics*, 88(2), 599–615. <https://doi.org/10.1007/s11192-011-0406-x>

Adamopoulos, P., Ghose, A., & Todri, V. (2018). The impact of user personality traits on word of mouth: Text-mining social media platforms. *Information Systems Research*, 29(3), 612–640. Scopus. <https://doi.org/10.1287/isre.2017.0768>

Adinolfi, F., De Rosa, M., & Trabalzi, F. (2011). Dedicated and generic marketing strategies The disconnection between geographical indications and consumer behavior in Italy. *BRITISH FOOD JOURNAL*, 113(2–3), 419–435. <https://doi.org/10.1108/00070701111116473>

Adler, N., Brudner, A., Gallotti, R., Privitera, F., & Ramasco, J. J. (2022). Does big data help answer big questions? The case of airport catchment areas & competition. *Transportation Research Part B: Methodological*, 166, 444–467. Scopus. <https://doi.org/10.1016/j.trb.2022.10.013>

Alikhani, M., Moghaddam, M. P., & Moazzen, F. (2022). Optimal demand response programs selection using CNN-LSTM algorithm with big data analysis of load curves. *IET Generation, Transmission and Distribution*, 16(24), 4980–5001. Scopus. <https://doi.org/10.1049/gtd2.12650>

Almansour, A. (2021). USING BIG DATA IN THE REAL ESTATE SECTOR. *ADVANCES AND APPLICATIONS IN STATISTICS*, 67(2), 207–224. <https://doi.org/10.17654/AS067020207>

Alyoubi, B. A. (2019). The impact of big data on electronic commerce in profit organisations in Saudi Arabia. *Research in World Economy*, 10(4), 106–115. Scopus. <https://doi.org/10.5430/rwe.v10n4p106>

Anderson, K., Burford, O., & Emmerton, L. (2016). App Chronic Disease Checklist: Protocol to Evaluate Mobile Apps for Chronic Disease Self-Management. *JMIR RESEARCH PROTOCOLS*, 5(4), e204. <https://doi.org/10.2196/resprot.6194>

Babiceanu, R. F., & Seker, R. (2016). Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook. *COMPUTERS IN INDUSTRY*, 81, 128–137. <https://doi.org/10.1016/j.compind.2016.02.004>

Barbado, R., Araque, O., & Iglesias, C. A. (2019). A framework for fake review detection in online consumer electronics retailers. *INFORMATION PROCESSING & MANAGEMENT*, 56(4), 1234–1244. <https://doi.org/10.1016/j.ipm.2019.03.002>

Barbera, G., Araujo, L., & Fernandes, S. (2023). The Value of Web Data Scraping: An Application to TripAdvisor. *Big Data and Cognitive Computing*, 7(3). Scopus. <https://doi.org/10.3390/bdcc7030121>

Barnes, S. J., Diaz, M., & Arnaboldi, M. (2021). Understanding panic buying during COVID-19: A text analytics approach. *Expert Systems with Applications*, 169. Scopus. <https://doi.org/10.1016/j.eswa.2020.114360>

Bello-Orgaz, G., Jung, J. J., & Camacho, D. (2016). Social big data: Recent achievements and new challenges. *INFORMATION FUSION*, 28, 45–59. <https://doi.org/10.1016/j.inffus.2015.08.005>

Biswas, S., Fole, A., Khare, N., & Agrawal, P. (2023). Enhancing correlated big data privacy using differential privacy and machine learning. *JOURNAL OF BIG DATA*, 10(1), 30. <https://doi.org/10.1186/s40537-023-00705-8>

Brandtner, P., Darbanian, F., Falatouri, T., & Udokwu, C. (2021). Impact of COVID-19 on the customer end of retail supply chains: A big data analysis of consumer satisfaction. *Sustainability (Switzerland)*, 13(3), 1–18. Scopus. <https://doi.org/10.3390/su13031464>

Buettner, R. (2017). Predicting user behavior in electronic markets based on personality-mining in large online social networks. *ELECTRONIC MARKETS*, 27(3), 247–265. <https://doi.org/10.1007/s12525-016-0228-z>

Chandra, S., & Verma, S. (2023). Big Data and Sustainable Consumption: A Review and Research Agenda. *VISION-THE JOURNAL OF BUSINESS PERSPECTIVE*, 27(1), 11–23. <https://doi.org/10.1177/09722629211022520>

Chang, S.-Y., Bodolica, V., Hsu, H.-H., & Lu, H.-P. (2023). What people talk about online and what they intend to do: Related perspectives from text mining and path analysis. *Eurasian Business Review*, 13(4), 931–956. Scopus. <https://doi.org/10.1007/s40821-022-00221-4>

Chen, C. M. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY*, 57(3), 359–377. <https://doi.org/10.1002/asi.20317>

Chen, M., & Xia, Z. (2015). Deduction of customer behaviour system under the background of big data. *Oxidation Communications*, 38(2A), 1142–1153. Scopus.

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. Scopus. <https://doi.org/10.1016/j.techfore.2017.06.034>

Clark, S. D., Shute, B., Jenneson, V., Rains, T., Birkin, M., & Morris, M. A. (2021). Dietary patterns derived from UK supermarket transaction data with nutrient and socioeconomic profiles. *Nutrients*, 13(5). Scopus. <https://doi.org/10.3390/nu13051481>

Chaudhary, K., Alam, M., Al-Rakhami, M. S., & Gumei, A. (2021). Machine learning-based mathematical modelling for prediction of social media consumer behavior using big data analytics. *JOURNAL OF BIG DATA*, 8(1), 73. <https://doi.org/10.1186/s40537-021-00466-2>

Chen, C. L. P., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *INFORMATION SCIENCES*, 275, 314–347. <https://doi.org/10.1016/j.ins.2014.01.015>

Dabrowska, A. (2011). Consumer behaviour in the market of catering services in selected countries of Central-Eastern Europe. *BRITISH FOOD JOURNAL*, 113(1), 96–108. <https://doi.org/10.1108/0007070111097367>

Daviet, R., Nave, G., & Wind, J. (2022). Genetic Data: Potential Uses and Misuses in Marketing. *Journal of Marketing*, 86(1), 7–26. Scopus. <https://doi.org/10.1177/0022242920980767>

de Luca, P., Pegan, G., & Gonzalo-Penela, C. (2019). Insights from a Google Keywords Analysis about Italian Wine in the US Market. *Micro and Macro Marketing*, 1, 93–116. Scopus. <https://doi.org/10.1431/93018>

Desarkar, A., & Das, A. (2017). Big-Data Analytics, Machine Learning Algorithms and Scalable/Parallel/Distributed Algorithms. In C. Bhatt, N. Dey, & A. S. Ashour (Eds.), *INTERNET OF THINGS AND BIG DATA TECHNOLOGIES FOR NEXT GENERATION HEALTHCARE* (Vol. 23, pp. 159–197). Springer International Publishing Ag. https://doi.org/10.1007/978-3-319-49736-5_8

Diem, A., & Wolter, S. C. (2013). The Use of Bibliometrics to Measure Research Performance in Education Sciences. *Research in Higher Education*, 54(1), 86–114. <https://doi.org/10.1007/s11162-012-9264-5>

Dinu, D., Stoica, I., & Radu, A. V. (2016). Studying the consumer behavior through big data. *Quality - Access to Success*, 17, 246–254. Scopus.

Dong, L. (2022). Analysis on Influencing Factors of Consumer Trust in E-Commerce Marketing of Green Agricultural Products Based on Big Data Analysis. *Mathematical Problems in Engineering*, 2022. Scopus. <https://doi.org/10.1155/2022/8221657>

Du, G., & Lin, Y. (2022). Brand connection and entry in the shopping mall ecological chain: Evidence from consumer behavior big data analysis based on two-sided markets. *Journal of Cleaner Production*, 364. Scopus. <https://doi.org/10.1016/j.jclepro.2022.132663>

Dubé, L., Labban, A., Moubarac, J.-C., Heslop, G., Ma, Y., & Paquet, C. (2014). A nutrition/health mindset on commercial Big Data and drivers of food demand in modern and traditional systems. *Annals of the New York Academy of Sciences*, 1331(1), 278–295. Scopus. <https://doi.org/10.1111/nyas.12595>

Dutta, C. B., & Das, D. K. (2017). What drives consumers' online information search behavior? Evidence from England. *JOURNAL OF RETAILING AND CONSUMER SERVICES*, 35, 36–45. <https://doi.org/10.1016/j.jretconser.2016.10.015>

Eck, N. van, & Waltman, L. (2009). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>

Ehsani, F., & Hosseini, M. (2023). Consumer Segmentation Based on Location and Timing Dimensions Using Big Data from Business-to-Customer Retailing Marketplaces. *Big Data. Scopus.* <https://doi.org/10.1089/big.2022.0307>

El Majdouli, M. A., Rbouh, I., Bougrine, S., El Benani, B., & El Imrani, A. A. (2016). Fireworks algorithm framework for Big Data optimization. *MEMETIC COMPUTING, 8(4), 333–347.* <https://doi.org/10.1007/s12293-016-0201-6>

Evangelin, M., & Vasantha, S. (2022). Mechanism of Big Data Analytics in Consumer Behavior on Online Shopping. *INTERNATIONAL JOURNAL OF EARLY CHILDHOOD SPECIAL EDUCATION, 14(3), 1938–1942.* <https://doi.org/10.9756/INT-JECSE/V14I3.226>

Fang, Y., Wang, H., Zhao, L., Yu, F., & Wang, C. (2020). Dynamic knowledge graph based fake-review detection. *APPLIED INTELLIGENCE, 50(12), 4281–4295.* <https://doi.org/10.1007/s10489-020-01761-w>

Félix, B. M., Tavares, E., & Cavalcante, N. W. F. (2018). Critical success factors for big data adoption in the virtual retail: Magazine luiza case study. *Revista Brasileira de Gestao de Negocios, 20(1), 112–126.* Scopus. <https://doi.org/10.7819/rbgn.v20i1.3627>

Fernández-Rovira, C., Álvarez Valdés, J., Molleví, G., & Nicolas-Sans, R. (2021). The digital transformation of business. Towards the datafication of the relationship with customers. *Technological Forecasting and Social Change, 162.* Scopus. <https://doi.org/10.1016/j.techfore.2020.120339>

Filieri, R., Lin, Z., Li, Y., Lu, X., & Yang, X. (2022). Customer Emotions in Service Robot Encounters: A Hybrid Machine-Human Intelligence Approach. *JOURNAL OF SERVICE RESEARCH, 25(4), 614–629.* <https://doi.org/10.1177/10946705221103937>

Filieri, R., Milone, F. L., Paolucci, E., & Raguseo, E. (2023). A big data analysis of COVID-19 impacts on Airbnbs' bookings behavior applying construal level and signaling theories. *International Journal of Hospitality Management, 111, 103461.* Scopus. <https://doi.org/10.1016/j.ijhm.2023.103461>

Fotopoulou, E., Zafeiropoulos, A., Terroso-Sáenz, F., Şimşek, U., González-Vidal, A., Tsoli, G., Gouvas, P., Liapis, P., Fensel, A., & Skarmeta, A. (2017). Providing personalized energy management and awareness services for energy efficiency in smart buildings. *Sensors (Switzerland), 17(9).* Scopus. <https://doi.org/10.3390/s17092054>

Gan, M., & Ouyang, Y. (2022). Study on Tourism Consumer Behavior Characteristics Based on Big Data Analysis. *FRONTIERS IN PSYCHOLOGY, 13.* <https://doi.org/10.3389/fpsyg.2022.876993>

Ghose, A., & Todri-Adamopoulos, V. (2016). TOWARD A DIGITAL ATTRIBUTION MODEL: MEASURING THE IMPACT OF DISPLAY ADVERTISING ON ONLINE CONSUMER BEHAVIOR. *MIS QUARTERLY, 40(4), 889–.* <https://doi.org/10.25300/MISQ/2016/40.4.05>

Giglio, S., Pantano, E., Bilotta, E., & Melewar, T. C. (2020). Branding luxury hotels: Evidence from the analysis of consumers' "big" visual data on TripAdvisor. *Journal of Business Research, 119, 495–501.* Scopus. <https://doi.org/10.1016/j.jbusres.2019.10.053>

Golmohammadi, A., Havakhor, T., Gauri, D., & Comprix, J. (2021). Complaint Publicization in Social Media. *JOURNAL OF MARKETING, 85(6), 1–23.* <https://doi.org/10.1177/00222429211002183>

Grandhi, B., Patwa, N., & Saleem, K. (2021). Data-driven marketing for growth and profitability. *EuroMed Journal of Business, 16(4), 381–398.* Scopus. <https://doi.org/10.1108/EMJB-09-2018-0054>

Green, M. A., Watson, A. W., Brunstrom, J. M., Corfe, B. M., Johnstone, A. M., Williams, E. A., & Stevenson, E. (2020). Comparing supermarket loyalty card data with traditional diet survey data for understanding how protein is purchased and consumed in older adults for the UK, 2014-16. *Nutrition Journal, 19(1).* Scopus. <https://doi.org/10.1186/s12937-020-00602-3>

Guo, L., & Zhang, D. (2019). EC-structure: Establishing consumption structure through mining E-commerce data to discover consumption upgrade. *Complexity, 2019.* Scopus. <https://doi.org/10.1155/2019/6543590>

Han, M. (2023). Consumer Behavior Analysis and Personalized Marketing Strategies for the Internet of Things. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao, 2023(Special issue E63), 367–376.* Scopus.

Han, S., & Lee, Y. (2022). Analysis of the impacts of social class and lifestyle on consumption of organic foods in South Korea. *Heliyon, 8(10).* Scopus. <https://doi.org/10.1016/j.heliyon.2022.e10998>

Hausladen, I., & Zipf, T. (2018). Competitive differentiation versus commoditisation: The role of big data in the european payments industry. *Journal of Payments Strategy and Systems, 12(3), 266–282.* Scopus.

Havranek, T., & Zeynalov, A. (2021). Forecasting tourist arrivals: Google Trends meets mixed-frequency data. *TOURISM ECONOMICS*, 27(1), 129–148. <https://doi.org/10.1177/1354816619879584>

Hofacker, C. F., Malthouse, E. C., & Sultan, F. (2016). Big Data and consumer behavior: Imminent opportunities. *Journal of Consumer Marketing*, 33(2), 89–97. Scopus. <https://doi.org/10.1108/JCM-04-2015-1399>

Hosseini, M., & Ghalamkari, A. (2018). Analysis Social Media Based Brand Communities and Consumer Behavior: A Netnographic Approach. *INTERNATIONAL JOURNAL OF E-BUSINESS RESEARCH*, 14(1), 37–53. <https://doi.org/10.4018/IJEBR.2018010103>

Ierkens, J. B., Fearnhead, P., & Roberts, G. (2019). The Zig-Zag Process and Super-Efficient Sampling for Bayesian Analysis of Big Data. *ANNALS OF STATISTICS*, 47(3), 1288–1320. <https://doi.org/10.1214/18-AOS1715>

Izang, A. A., Goga, N., Kuyoro, S. O., Alao, O. D., Omotunde, A. A., & Adio, A. K. (2019). Scalable data analytics market basket model for transactional data streams. *International Journal of Advanced Computer Science and Applications*, 10(10), 61–68. Scopus. <https://doi.org/10.14569/ijacsa.2019.0101010>

Jackson, I., & Ivanov, D. (2023). A beautiful shock? Exploring the impact of pandemic shocks on the accuracy of AI forecasting in the beauty care industry. *TRANSPORTATION RESEARCH PART E-LOGISTICS AND TRANSPORTATION REVIEW*, 180, 103360. <https://doi.org/10.1016/j.tre.2023.103360>

Jahns, J. (2023). The future of marketing: How cultural understanding contributes to the success of brand positioning and campaigns. *Journal of Digital and Social Media Marketing*, 11(3), 225–235. Scopus.

Jiménez-Marín, G., Sanz-Marcos, P., Medina, I. G., & Coelho, P. M. F. (2020). How Big Data Collected Via Point of Sale Devices in Textile Stores in Spain Resulted in Effective Online Advertising Targeting. *International Journal of Interactive Mobile Technologies*, 14(13), 65–77. Scopus. <https://doi.org/10.3991/ijim.v14i13.14359>

Jung, J., Yeo, S., Lee, Y., Moon, S., & Lee, D.-J. (2021). Factors affecting consumers' preferences for electric vehicle: A Korean case. *RESEARCH IN TRANSPORTATION BUSINESS AND MANAGEMENT*, 41, 100666. <https://doi.org/10.1016/j.rtbm.2021.100666>

Kanavos, A., Iakovou, S. A., Sioutas, S., & Tampakas, V. (2018). Large scale product recommendation of supermarket ware based on customer behaviour analysis. *Big Data and Cognitive Computing*, 2(2), 1–19. Scopus. <https://doi.org/10.3390/bdcc2020011>

Kim, D., Woo, J., Shin, J., Lee, J., & Kim, Y. (2019). Can search engine data improve accuracy of demand forecasting for new products? Evidence from automotive market. *Industrial Management and Data Systems*, 119(5), 1089–1103. Scopus. <https://doi.org/10.1108/IMDS-08-2018-0347>

Kim, W.-H., Park, E., & Kim, S.-B. (2023). Understanding the role of firm-generated content by hotel segment: The case of Twitter. *Current Issues in Tourism*, 26(1), 122–136. Scopus. <https://doi.org/10.1080/13683500.2021.2003759>

Kiran, P., & Vasantha, S. (2016). Analysing the role of user generated content on consumer purchase intention in the new era of social media and big data. *Indian Journal of Science and Technology*, 9(43). Scopus. <https://doi.org/10.17485/ijst/2016/v9i43/104754>

Labrecque, L. I. (2014). Fostering Consumer-Brand Relationships in Social Media Environments: The Role of Parasocial Interaction. *JOURNAL OF INTERACTIVE MARKETING*, 28(2), 134–148. <https://doi.org/10.1016/j.intmar.2013.12.003>

Lan, M. (2023). Big Data and Cloud Computing-Integrated Tourism Decision-Making in Smart Logistics Technologies. *INTERNATIONAL JOURNAL OF E-COLLaborATION*, 19(7). <https://doi.org/10.4018/ijec.316880>

Lee, C. K. H., Tse, Y. K., Zhang, M., & Ma, J. (2020). Analysing online reviews to investigate customer behaviour in the sharing economy: The case of Airbnb. *Information Technology and People*, 33(3), 945–961. Scopus. <https://doi.org/10.1108/ITP-10-2018-0475>

Lee, J., Jung, O., Lee, Y., Kim, O., & Park, C. (2021). A comparison and interpretation of machine learning algorithm for the prediction of online purchase conversion. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(5), 1472–1491. Scopus. <https://doi.org/10.3390/jtaer16050083>

Lee, Y., & Kim, D.-Y. (2020). The decision tree for longer-stay hotel guest: The relationship between hotel booking determinants and geographical distance. *International Journal of Contemporary Hospitality Management*, 33(6), 2264–2282. Scopus. <https://doi.org/10.1108/IJCHM-06-2020-0594>

Li, J., & Hu, Q. (2021). Using culturomics and social media data to characterize wildlife consumption. *Conservation Biology*, 35(2), 452–459. Scopus. <https://doi.org/10.1111/cobi.13703>

Li, L. (2023). Analysis of e-commerce customers' shopping behavior based on data mining and machine learning. *Soft Computing*. Scopus. <https://doi.org/10.1007/s00500-023-08903-5>

Li, R., Xu, X., & Dong, S. (2022). Construction of Precision Sales Model for Luxury Market Based on Machine Learning. *Mobile Information Systems*, 2022. Scopus. <https://doi.org/10.1155/2022/7920628>

Li, X. T., & Feng, F. (2018). Enterprise customer relationship management based on big data mining. *Latin American Applied Research*, 48(3), 163–168. Scopus.

Liu, H. (2021). Big data precision marketing and consumer behavior analysis based on fuzzy clustering and PCA model. *Journal of Intelligent and Fuzzy Systems*, 40(4), 6529–6539. Scopus. <https://doi.org/10.3233/JIFS-189491>

Liu, H., Huang, Y., Wang, Z., Liu, K., Hu, X., & Wang, W. (2019). Personality or Value: A Comparative Study of Psychographic Segmentation Based on an Online Review Enhanced Recommender System. *APPLIED SCIENCES-BASEL*, 9(10). <https://doi.org/10.3390/app9101992>

Liu, X., Singh, P. V., & Srinivasan, K. (2016). A structured analysis of unstructured big data by leveraging cloud computing. *Marketing Science*, 35(3), 363–388. Scopus. <https://doi.org/10.1287/mksc.2015.0972>

Liu, X., & Zhao, H. (2021). Dairy brand loyalty measurement model based on machine learning clustering algorithm. *Journal of Intelligent and Fuzzy Systems*, 40(4), 7601–7612. Scopus. <https://doi.org/10.3233/JIFS-189580>

Liu, Y. (2016). Design and implementation of hadoop-based customer marketing big data processing system. *International Journal of Database Theory and Application*, 9(12), 331–340. Scopus. <https://doi.org/10.14257/ijdta.2016.9.12.29>

Liu, Y., Francis, A., Hollauer, C., Lawson, M., Shaikh, O., Cotsman, A., Bhardwaj, K., Banboukian, A., Li, M., Webb, A., & Asensio, O. (2023). Reliability of electric vehicle charging infrastructure: A cross-lingual deep learning approach. *COMMUNICATIONS IN TRANSPORTATION RESEARCH*, 3. <https://doi.org/10.1016/j.commtr.2023.100095>

Lv, H. (2022). Intelligent e-commerce framework for consumer behavior analysis using big data Analytics. *ADVANCES IN DATA SCIENCE AND ADAPTIVE ANALYSIS*, 14(03N04). <https://doi.org/10.1142/S2424922X22500073>

Lv, H. (2023). E-commerce consumer behavior analysis based on big data. *Journal of Computational Methods in Sciences and Engineering*, 23(2), 651–661. Scopus. <https://doi.org/10.3233/JCM-226628>

Lv, H., Shi, S., & Gursoy, D. (2022). A look back and a leap forward: A review and synthesis of big data and artificial intelligence literature in hospitality and tourism. *JOURNAL OF HOSPITALITY MARKETING & MANAGEMENT*, 31(2), 145–175. <https://doi.org/10.1080/19368623.2021.1937434>

Maeng, Y., Lee, C. C., & Yun, H. (2023). Understanding Antecedents That Affect Customer Evaluations of Head-Mounted Display VR Devices through Text Mining and Deep Neural Network. *Journal of Theoretical and Applied Electronic Commerce Research*, 18(3), 1238–1256. Scopus. <https://doi.org/10.3390/jtaer18030063>

Makov, T., & Fitzpatrick, C. (2021). Is repairability enough? Big data insights into smartphone obsolescence and consumer interest in repair. *Journal of Cleaner Production*, 313. Scopus. <https://doi.org/10.1016/j.jclepro.2021.127561>

Mariani, M., Borghi, M., & Kazakov, S. (2019). The role of language in the online evaluation of hospitality service encounters: An empirical study. *INTERNATIONAL JOURNAL OF HOSPITALITY MANAGEMENT*, 78, 50–58. <https://doi.org/10.1016/j.ijhm.2018.11.012>

Mariani, M., & Predvoditeleva, M. (2019). How do online reviewers' cultural traits and perceived experience influence hotel online ratings? An empirical analysis of the Muscovite hotel sector. *INTERNATIONAL JOURNAL OF CONTEMPORARY HOSPITALITY MANAGEMENT*, 31(12), 4543–4573. <https://doi.org/10.1108/IJCHM-11-2018-0927>

Martens, D., Provost, F., Clark, J., & de Fortuny, E. J. (2016). Mining massive fine-grained behavior data to improve predictive analytics. *MIS Quarterly: Management Information Systems*, 40(4), 869–888. Scopus. <https://doi.org/10.25300/MISQ/2016/40.4.04>

Martin, K. D., & Murphy, P. E. (2017). The role of data privacy in marketing. *JOURNAL OF THE ACADEMY OF MARKETING SCIENCE*, 45(2), 135–155. <https://doi.org/10.1007/s11747-016-0495-4>

Mayr, P., & Scharnhorst, A. (2015). Scientometrics and information retrieval: Weak-links revitalized. *Scientometrics*, 102(3), 2193–2199. <https://doi.org/10.1007/s11192-014-1484-3>

Mendieta-Aragón, A., & Garín-Muñoz, T. (2023). Consumer behaviour in e-Tourism: Exploring new applications of machine learning in tourism studies. *Investigaciones Turísticas*, 26, 350–374. Scopus. <https://doi.org/10.14198/INTURI24629>

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & the PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine*, 151(4), 264–269. <https://doi.org/10.7326/0003-4819-151-4-200908180-00135>

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1), 1. <https://doi.org/10.1186/2046-4053-4-1>

Molitor, D., Daurer, S., Spann, M., & Manchanda, P. (2023). Digitizing local search: An empirical analysis of mobile search behavior in offline shopping. *Decision Support Systems*, 174. Scopus. <https://doi.org/10.1016/j.dss.2023.114018>

Moon, J.-Y., & Hong, J.-W. (2018). The differential effects of online content on healthcare adoption: Hierarchical modelling. *Journal of Theoretical and Applied Information Technology*, 96(6), 1722–1731. Scopus.

Morris, M. A., Wilkins, E. L., Galazoula, M., Clark, S. D., & Birkin, M. (2020). Assessing diet in a university student population: A longitudinal food card transaction data approach. *British Journal of Nutrition*, 123(12), 1406–1414. Scopus. <https://doi.org/10.1017/S0007114520000823>

Moschis, G. P. (2012). Consumer Behavior in Later Life: Current Knowledge, Issues, and New Directions for Research. *PSYCHOLOGY & MARKETING*, 29(2), 57–75. <https://doi.org/10.1002/mar.20504>

Mu, W. (2019). A Big Data-based Prediction Model for Purchase Decisions of Consumers on Cross-border E-commerce Platforms. *Journal European Des Systèmes Automatisés*, 52(4), 363–368. Scopus. <https://doi.org/10.18280/jesa.520405>

Naraine, M. L., O'Reilly, N., Levallet, N., & Wanless, L. (2020). If you build it, will they log on? Wi-Fi usage and behavior while attending National Basketball Association games. *Sport, Business and Management: An International Journal*, 10(2), 207–226. Scopus. <https://doi.org/10.1108/SBM-02-2019-0016>

Nie, Y., & Han, X. (2019). Research on consumers' protection in advantageous operation of big data brokers. *Cluster Computing*, 22, 8387–8400. Scopus. <https://doi.org/10.1007/s10586-018-1837-4>

Nilashi, M., Abumaloh, R. A., Alghamdi, A., Minaei-Bidgoli, B., Alsulami, A. A., Thanoon, M., Asadi, S., & Samad, S. (2021). What is the impact of service quality on customers' satisfaction during COVID-19 outbreak? New findings from online reviews analysis. *Telematics and Informatics*, 64. Scopus. <https://doi.org/10.1016/j.tele.2021.101693>

Oh, J., Connerton, T. P., & Kim, H.-J. (2019). The rediscovery of brand experience dimensions with big data analysis: Building for a sustainable brand. *Sustainability (Switzerland)*, 11(19). Scopus. <https://doi.org/10.3390/su11195438>

Oprea, S.-V., & Bâra, A. (2020). Setting the Time-of-Use Tariff Rates with NoSQL and Machine Learning to a Sustainable Environment. *IEEE Access*, 8, 25521–25530. Scopus. <https://doi.org/10.1109/ACCESS.2020.2969728>

Oprea, S.-V., Bâra, A., Tudorică, B. G., Călinou, M. I., & Botezatu, M. A. (2021). Insights into demand-side management with big data analytics in electricity consumers' behaviour. *Computers and Electrical Engineering*, 89. Scopus. <https://doi.org/10.1016/j.compeleceng.2020.106902>

Ozturkcan, S., Kasap, N., Tanaltay, A., & Ozdinc, M. (2019). Analysis of tweets about football: 2013 and 2018 leagues in Turkey. *Behaviour and Information Technology*, 38(9), 887–899. Scopus. <https://doi.org/10.1080/0144929X.2019.1583284>

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. <https://doi.org/10.1136/bmj.n71>

Pai, C.-S., & Chen, S.-L. (2023). Mystery of Big Data: A Study of Consumer Decision-Making Behavior on E-Commerce Websites. *Engineering Proceedings*, 38(1). Scopus. <https://doi.org/10.3390/engproc2023038029>

Pang, W., Ko, J., Kim, S., & Ko, E. (2022). Impact of COVID-19 pandemic upon fashion consumer behavior: Focus on mass and luxury products. *ASIA PACIFIC JOURNAL OF MARKETING AND LOGISTICS*, 34(10), 2149–2164. <https://doi.org/10.1108/APJML-03-2021-0189>

Pantano, E., Giglio, S., & Dennis, C. (2019). Making sense of consumers' tweets: Sentiment outcomes for fast fashion retailers through Big Data analytics. *International Journal of Retail and Distribution Management*, 47(9), 915–927. Scopus. <https://doi.org/10.1108/IJRD-07-2018-0127>

Pindado, E., & Barrena, R. (2021). Using Twitter to explore consumers' sentiments and their social representations towards new food trends. *British Food Journal*, 123(3), 1060–1082. Scopus. <https://doi.org/10.1108/BFJ-03-2020-0192>

Praveen Kumar, Y., Suguna, R., Ghosh, S., & Neha, K. (2019). Semantic web mining in retail management system using ANN. *International Journal of Recent Technology and Engineering*, 8(2 Special Issue 11), 3547–3554. Scopus. <https://doi.org/10.35940/ijrte.B1439.0982S1119>

Prentice, C., Chen, J., & Stantic, B. (2020). Timed intervention in COVID-19 and panic buying. *Journal of Retailing and Consumer Services*, 57. Scopus. <https://doi.org/10.1016/j.jretconser.2020.102203>

Ramos, C. M. Q., Martins, D. J., Serra, F., Lam, R., Cardoso, P. J. S., Correia, M. B., & Rodrigues, J. M. F. (2017). Framework for a hospitality big data warehouse: The implementation of an efficient hospitality business intelligence system. *International Journal of Information Systems in the Service Sector*, 9(2), 27–45. Scopus. <https://doi.org/10.4018/IJISSS.2017040102>

Rendon, E., Alejo, R., Castorena, C., Isidro-Ortega, F. J., & Granda-Gutierrez, E. E. (2020). Data Sampling Methods to Deal With the Big Data Multi-Class Imbalance Problem. *APPLIED SCIENCES-BASEL*, 10(4), 1276. <https://doi.org/10.3390/app10041276>

Ryu, G.-A., Nasridinov, A., Rah, H., & Yoo, K.-H. (2020). Forecasts of the amount purchase pork meat by using structured and unstructured big data. *Agriculture (Switzerland)*, 10(1). Scopus. <https://doi.org/10.3390/agriculture10010021>

Sakas, D., Kamperos, I., & Reklitis, P. (2021). Estimating Risk Perception Effects on Courier Companies' Online Customer Behavior during a Crisis, Using Crowdsourced Data. *SUSTAINABILITY*, 13(22). <https://doi.org/10.3390/su132212725>

Sarasquete, N. (2017). A common data representation model for customer behavior tracking. *REVISTA ICONO 14-REVISTA CIENTIFICA DE COMUNICACION Y TECNOLOGIAS*, 15(2), 55–91. <https://doi.org/10.7195/ri14.1078v15i2.1078>

Sarkis-Onofre, R., Catalá-López, F., Aromataris, E., & Lockwood, C. (2021). How to properly use the PRISMA Statement. *Systematic Reviews*, 10(1), 117. <https://doi.org/10.1186/s13643-021-01671-z>

Senavirathne, N., & Torra, V. (2020). On the Role of Data Anonymization in Machine Learning Privacy. In G. J. Wang, R. Ko, M. Z. A. Bhuiyan, & Y. Pan (Eds.), *2020 IEEE 19TH INTERNATIONAL CONFERENCE ON TRUST, SECURITY AND PRIVACY IN COMPUTING AND COMMUNICATIONS (TRUSTCOM 2020)* (pp. 664–675). IEEE Computer Soc. <https://doi.org/10.1109/TrustCom50675.2020.00093>

Seo, D., & Yoo, Y. (2023). Improving Shopping Mall Revenue by Real-Time Customized Digital Coupon Issuance. *IEEE Access*, 11, 7924–7932. Scopus. <https://doi.org/10.1109/ACCESS.2023.3239425>

Serrano, L., Ariza-Montes, A., Nader, M., Sianes, A., & Law, R. (2021). Exploring preferences and sustainable attitudes of Airbnb green users in the review comments and ratings: A text mining approach. *Journal of Sustainable Tourism*, 29(7), 1134–1152. Scopus. <https://doi.org/10.1080/09669582.2020.1838529>

Shang, P., & Li, T. (2017). Consumers repurchasing behavior research based on big data environment. *Agro Food Industry Hi-Tech*, 28(3), 1206–1210. Scopus.

Shibuya, Y., & Tanaka, H. (2018). A Statistical Analysis Between Consumer Behavior and a Social Network Service: A Case Study of Used-Car Demand Following the Great East Japan Earthquake and Tsunami of 2011. *REVIEW OF SOCIONETWORK STRATEGIES*, 12(2), 205–236. <https://doi.org/10.1007/s12626-018-0025-6>

Shokouhyar, S., Dehkhodaei, A., & Amiri, B. (2022). Toward customer-centric mobile phone reverse logistics: Using the DEMATEL approach and social media data. *Kybernetes*, 51(11), 3236–3279. Scopus. <https://doi.org/10.1108/K-11-2020-0831>

Silva, E. S., Hassani, H., & Madsen, D. Ø. (2020). Big Data in fashion: Transforming the retail sector. *Journal of Business Strategy*, 41(4), 21–27. Scopus. <https://doi.org/10.1108/JBS-04-2019-0062>

Silva, E. S., Hassani, H., Madsen, D. Ø., & Gee, L. (2019). Googling fashion: Forecasting fashion consumer behaviour using Google Trends. *Social Sciences*, 8(4). Scopus. <https://doi.org/10.3390/socsci8040111>

Šimović, P. P., Chen, C. Y. T., & Sun, E. W. (2023). Classifying the Variety of Customers' Online Engagement for Churn Prediction with a Mixed-Penalty Logistic Regression. *Computational Economics*, 61(1), 451–485. Scopus. <https://doi.org/10.1007/s10614-022-10275-1>

Singh, A., & Glińska-Neweś, A. (2022). Modeling the public attitude towards organic foods: A big data and text mining approach. *Journal of Big Data*, 9(1). Scopus. <https://doi.org/10.1186/s40537-021-00551-6>

Singh, N., & Singh, A. K. (2018). Data Privacy Protection Mechanisms in Cloud. *DATA SCIENCE AND ENGINEERING*, 3(1), 24–39. <https://doi.org/10.1007/s41019-017-0046-0>

Singh, S., & Yassine, A. (2019). Mining energy consumption behavior patterns for households in smart grid. *IEEE Transactions on Emerging Topics in Computing*, 7(3), 404–419. Scopus. <https://doi.org/10.1109/TETC.2017.2692098>

Soria-Comas, J., & Domingo-Ferrer, J. (2016). Big Data Privacy: Challenges to Privacy Principles and Models. *DATA SCIENCE AND ENGINEERING*, 1(1), 21–28. <https://doi.org/10.1007/s41019-015-0001-x>

Sun, M., & Zhao, J. (2022). Behavioral Patterns beyond Posting Negative Reviews Online: An Empirical View. *Journal of Theoretical and Applied Electronic Commerce Research*, 17(3), 949–983. Scopus. <https://doi.org/10.3390/jtaer17030049>

Thu, H. P., The, N. N., Thi, T. H. P., & Nam, T. N. (2019). Evaluating the purchase behaviour of organic food by young consumers in an emerging market economy. *JOURNAL OF STRATEGIC MARKETING*, 27(6), 540–556. <https://doi.org/10.1080/0965254X.2018.1447984>

Tian, J., Zhang, Y., & Zhang, C. (2018). Predicting consumer variety-seeking through weather data analytics. *Electronic Commerce Research and Applications*, 28, 194–207. Scopus. <https://doi.org/10.1016/j.elerap.2018.02.001>

Tian, Y. (2022). An Effective Model for Consumer Need Prediction Using Big Data Analytics. *JOURNAL OF INTERCONNECTION NETWORKS*, 22(SUPP02), 2143008. <https://doi.org/10.1142/S0219265921430088>

Tiantian, A. (2022). Data mining analysis method of consumer behaviour characteristics based on social media big data. *International Journal of Web Based Communities*, 18(3–4), 224–237. Scopus. <https://doi.org/10.1504/IJWBC.2022.125492>

Tupikovskaja-Omovie, Z., & Tyler, D. (2021). Eye tracking technology to audit google analytics: Analysing digital consumer shopping journey in fashion m-retail. *International Journal of Information Management*, 59. Scopus. <https://doi.org/10.1016/j.ijinfomgt.2020.102294>

Udugama, I. A., Kelton, W., & Bayer, C. (2023). Digital twins in food processing: A conceptual approach to developing multi-layer digital models. *Digital Chemical Engineering*, 7. Scopus. <https://doi.org/10.1016/j.dche.2023.100087>

Upadhyay, U., Kumar, A., Sharma, G., Sharma, S., Arya, V., Panigrahi, P. K., & Gupta, B. B. (2024). A systematic data-driven approach for targeted marketing in enterprise information system. *ENTERPRISE INFORMATION SYSTEMS*. <https://doi.org/10.1080/17517575.2024.2356770>

Urkup, C., Bozkaya, B., & Sibel Salman, F. (2018). Customer mobility signatures and financial indicators as predictors in product recommendation. *PLoS ONE*, 13(7). Scopus. <https://doi.org/10.1371/journal.pone.0201197>

Ushakova, A., & Jankin Mikhaylov, S. (2020). Big data to the rescue? Challenges in analysing granular household electricity consumption in the United Kingdom. *Energy Research and Social Science*, 64. Scopus. <https://doi.org/10.1016/j.erss.2020.101428>

Venkatrama, S. (2017). A proposed business intelligent framework for recommender systems. *Informatics*, 4(4). Scopus. <https://doi.org/10.3390/informatics4040040>

Vepsäläinen, H., Nevalainen, J., Kinnunen, S., Itkonen, S. T., Meinilä, J., Männistö, S., Uusitalo, L., Fogelholm, M., & Erkkola, M. (2022). Do we eat what we buy? Relative validity of grocery purchase data as an indicator of food consumption in the LoCard study. *British Journal of Nutrition*, 128(9), 1780–1788. Scopus. <https://doi.org/10.1017/S0007114521004177>

Volkova, E., & Karpushkin, G. (2023). Marketing communications in luxury fashion retail in the era of big data. *Electronic Commerce Research*. Scopus. <https://doi.org/10.1007/s10660-023-09732-2>

Wang, B., Shen, C., Cai, Y., Dai, L., Gai, S., & Liu, D. (2023). Consumer culture in traditional food market: The influence of Chinese consumers to the cultural construction of Chinese barbecue. *Food Control*, 143. Scopus. <https://doi.org/10.1016/j.foodcont.2022.109311>

Wang, J., & Zhang, Y. (2021). Using cloud computing platform of 6G IoT in e-commerce personalized recommendation. *INTERNATIONAL JOURNAL OF SYSTEM ASSURANCE ENGINEERING AND MANAGEMENT*, 12(4), 654–666. <https://doi.org/10.1007/s13198-021-01059-1>

Wang, Y. (2022). Big Data Mining Method of Marketing Management Based on Deep Trust Network Model. *International Journal of Circuits, Systems and Signal Processing*, 16, 578–584. Scopus. <https://doi.org/10.46300/9106.2022.16.72>

Wei, W., Sivaparthipan, C. B., & Kumar, P. M. (2022). Online shopping behavior analysis for smart business using big data analytics and blockchain security. *International Journal of Modeling, Simulation, and Scientific Computing*, 13(4). Scopus. <https://doi.org/10.1142/S1793962322500532>

Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *TECHNOLOGICAL FORECASTING AND SOCIAL CHANGE*, 126, 3–13. <https://doi.org/10.1016/j.techfore.2015.12.019>

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

Xiao, B., & Piao, G. (2022). Analysis of Influencing Factors and Enterprise Strategy of Online Consumer Behavior Decision Based on Association Rules and Mobile Computing. *Wireless Communications and Mobile Computing*, 2022. Scopus. <https://doi.org/10.1155/2022/6849017>

Xie, T. (2023). Artificial intelligence and automatic recognition application in B2C e-commerce platform consumer behavior recognition. *Soft Computing*, 27(11), 7627–7637. Scopus. <https://doi.org/10.1007/s00500-023-08147-3>

Xu, M., & Chen, P. (2023). Big data analysis based on the correlation between live-streaming with goods, perceived value and consumer repurchase. *Applied Mathematics and Nonlinear Sciences*. Scopus. <https://doi.org/10.2478/amns.2023.1.00134>

Xue, L. (2023). Cluster analysis-based big data mining method of e-commerce consumer behaviour. *International Journal of Web Based Communities*, 19(1), 53–63. Scopus. <https://doi.org/10.1504/IJWBC.2023.10048319>

Yan, Y., Huang, C., Wang, Q., & Hu, B. (2020). Data mining of customer choice behavior in internet of things within relationship network. *International Journal of Information Management*, 50, 566–574. Scopus. <https://doi.org/10.1016/j.ijinfomgt.2018.11.013>

Yang, J., Min, D., & Kim, J. (2020). The use of big data and its effects in a diffusion forecasting model for Korean reverse mortgage subscribers. *Sustainability (Switzerland)*, 12(3). Scopus. <https://doi.org/10.3390/su12030979>

Ye, D., Muthu, B., & Kumar, P. (2022). Identifying Buying Patterns From Consumer Purchase History Using Big Data and Cloud Computing. *INTERNATIONAL JOURNAL OF DISTRIBUTED SYSTEMS AND TECHNOLOGIES*, 13(7). <https://doi.org/10.4018/IJDST.307957>

Ye, Y., Lu, X., & Lu, T. (2022). Examining the spillover effect of sustainable consumption on microloan repayment: A big data-based research. *Information and Management*, 59(5). Scopus. <https://doi.org/10.1016/j.im.2020.103288>

Zhang, J., Wu, J., & Gao, C. (2022). Consumption Behavior Analysis of E-commerce Users Based on K-means Algorithm. *Journal of Network Intelligence*, 7(4), 935–942. Scopus.

Zhang, L., Priestley, J., Demaio, J., Ni, S., & Tian, X. (2021). Measuring Customer Similarity and Identifying Cross-Selling Products by Community Detection. *Big Data*, 9(2), 132–143. Scopus. <https://doi.org/10.1089/big.2020.0044>

Zhang, W., & He, Y. (2019). Optimal policies for new and green remanufactured short-life-cycle products considering consumer behavior. *JOURNAL OF CLEANER PRODUCTION*, 214, 483–505. <https://doi.org/10.1016/j.jclepro.2018.12.213>

Zhang, W., & Wang, M. (2021). An improved deep forest model for prediction of e-commerce consumers' repurchase behavior. *PLOS ONE*, 16(9). <https://doi.org/10.1371/journal.pone.0255906>

Zhao, X., Lu, T., & Dai, Y. (2021). Individual Driver Crash Risk Classification Based on IoV Data and Offline Consumer Behavior Data. *Mobile Information Systems*, 2021. Scopus. <https://doi.org/10.1155/2021/6784026>

Zhou, F., Lim, M. K., He, Y., & Pratap, S. (2020). What attracts vehicle consumers' buying: A Saaty scale-based VIKOR (SSC-VIKOR) approach from after-sales textual perspective? *Industrial Management and Data Systems*, 120(1), 57–78. Scopus. <https://doi.org/10.1108/IMDS-01-2019-0034>

Zhukovskiy, Y. L., Kovalchuk, M. S., Batueva, D. E., & Senchilo, N. D. (2021). Development of an algorithm for regulating the load schedule of educational institutions based on the forecast of electric consumption within the framework of application of the demand response. *Sustainability (Switzerland)*, 13(24). Scopus. <https://doi.org/10.3390/su132413801>

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.