
Volume 6, Issue 7, 2022
Peer-reviewed, open-access journal dedicated to publishing high-quality original research articles, literature reviews, case studies, and theoretical papers that contribute to the understanding of human behavior and social phenomena.
https://studies.eigenpub.com/index.php/jhbs

# Geospatial Big Data and the Built Environment: Applications in Urban Planning and Infrastructure Management

**Santiago Ramirez**
Department of Geomatics Engineering, Universidad Nacional de Colombia, Colombia

**Khurshed Iqbal**
Department of Management Sciences, University College of Zhob(UCoZ)

## ABSTRACT

The exponential growth of geospatial big data stemming from a plethora of sources, including smartphones, sensors, and satellite imagery, has ushered in a new era of opportunities and challenges for the fields of urban planning and infrastructure management. This paper offers a comprehensive review of the applications of geospatial big data in these domains, shedding light on the myriad ways it has redefined urban analytics. Notably, it delves into the transformational potential of geospatial big data in optimizing land use, enhancing transportation systems, and streamlining resource allocation within urban landscapes. Furthermore, it scrutinizes the employment of data mining techniques to harness valuable insights from the wealth of citizen-generated geospatial data, illuminating the path toward more informed urban development. Despite the evident promise, this paper does not shy away from addressing the formidable challenges that accompany the adoption of geospatial big data in built environment disciplines. Issues concerning data quality, privacy, efficient storage, and the development of requisite analytical capabilities are all thoroughly examined, providing a holistic view of the landscape. In summation, geospatial big data emerges as a cornerstone for evidence-based decision-making in the realm of smart urban development. However, its successful implementation necessitates a strategic approach to data governance and a concerted effort to foster the analytical skills of professionals in the built environment, equipping them with the expertise required to navigate this dynamic and data-rich terrain.

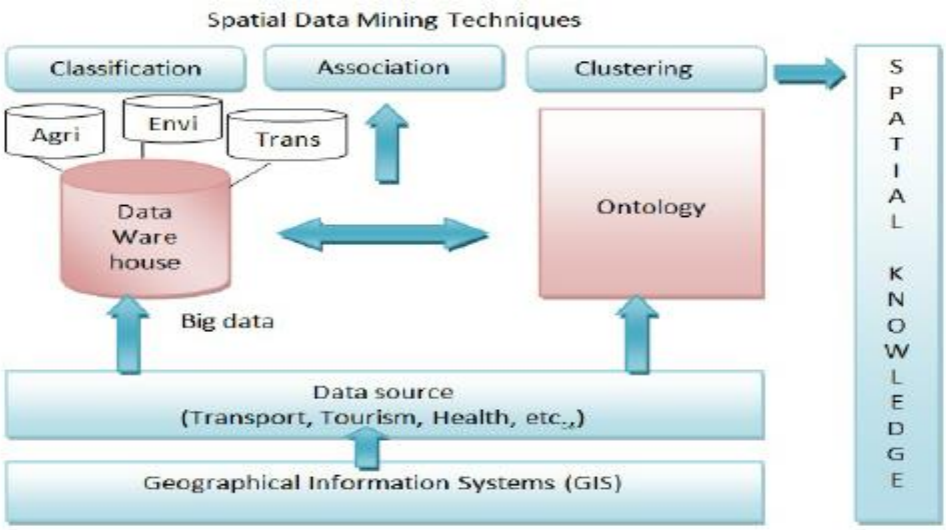
**Keywords**: *Geospatial big data, urban planning, infrastructure management, data mining, data quality, privacy*

## INTRODUCTION

The built environment, comprising man-made structures and landscapes, plays a pivotal role in shaping urban spaces. Urban planning and infrastructure management heavily rely on data and analytical tools to optimize the functioning and maintenance of these environments. Geospatial data, which encompasses information about the location and spatial relationships of elements within this environment, has long been integral to various built environment disciplines. However, the advent of digital devices and sensors has brought about an unprecedented surge in geospatial 'big data', presenting both opportunities and challenges for the field. Geospatial big data originates from diverse sources, including satellite imagery, GPS, social media check-ins, mobile devices, and various sensors embedded in urban infrastructure. These sources generate vast volumes of data with varying characteristics, such as velocity, variety, veracity, and volume, hence the term 'big data'. Analytical techniques such as Geographic Information Systems (GIS), spatial analysis, machine learning, and data mining have become essential in making sense of this wealth of information [1].

These tools offer the means to extract valuable insights, understand spatial patterns, and derive informed decisions crucial for effective urban planning and infrastructure management. The significance of geospatial big data in these realms cannot be overstated. In land use planning, it empowers authorities to make informed decisions about zoning, development, and allocation of resources. Transportation systems benefit from real-time data to optimize traffic flow and develop smarter mobility solutions. Resource allocation in urban areas, including emergency services distribution and environmental conservation, is enhanced by leveraging geospatial data [2]. Furthermore, citizen engagement and participation in planning processes have been revolutionized through interactive mapping and crowd-sourced data, fostering a more inclusive approach to decision-making [3]. Several impactful case studies and examples highlight the transformational potential of geospatial big data. For instance, the City of Chicago's Array of Things project collects real-time data on environmental factors, traffic, and other urban elements, enabling proactive decision-making. Similarly, in Los Angeles, the Mobility Data Specification initiative has revolutionized the transportation landscape by opening access to data collected by shared mobility operators, transforming how policymakers understand and plan for urban mobility needs [4].

Figure 1.




However, amidst the promise and potential, challenges persist. Data quality issues, including accuracy and completeness, remain a concern. Privacy concerns surrounding the collection and utilization of geospatial data raise ethical and legal considerations that require careful navigation [5]. The storage, processing, and accessibility of vast amounts of geospatial information pose logistical challenges. Additionally, there is a growing need for professionals in the built environment sector to acquire and hone the analytical skills necessary to derive meaningful insights from this deluge of data [6].

## Background

**Defining Geospatial Big Data:** Big data, often described by the '3 Vs' framework introduced by Kitchin (2014), encapsulates an immense and intricate landscape of data, posing unique challenges to traditional data processing tools. These three primary dimensions—volume, velocity, and variety—provide an insightful perspective on the vastness and complexity of big data. Volume pertains to the sheer magnitude of data being generated and collected, reaching unprecedented scales in the modern era. The proliferation of data from sources such as social media, sensors, and online transactions has contributed to this data deluge [7]. Velocity underscores the dynamic and real-time nature of data generation. In today's hyperconnected world, data is produced and updated at a remarkable pace, necessitating agile methods for its handling. Variety encompasses the diverse data types, formats, and sources that collectively constitute big data. Ranging from structured data in databases to unstructured text, images, and multimedia, big data embodies a rich tapestry of information.





In addition to the '3 Vs,' Laney (2001) introduced two supplementary dimensions that are of particular relevance in the context of big data. Veracity underscores the inherent uncertainty and inconsistency in data quality. Given the multitude of data sources and the potential for errors, ensuring data accuracy and reliability becomes a significant challenge. Value is another pivotal dimension, emphasizing the ultimate goal of deriving meaningful insights and utility from big data. It is imperative that the vast reservoirs of data translate into actionable information and tangible benefits [8]. Geospatial big data, on the other hand, is a distinct subset within the big data landscape, defined by Kitchin (2013) as extensive datasets that are continually updated and incorporate a geographical or location component. This unique category of data integrates spatial, temporal, or other geographic attributes, such as coordinates and location tags, providing an indispensable geographical context to the information. The key defining characteristic of geospatial big data is its spatial dimension, which adds a layer of complexity and richness to the dataset [9].

Within the built environment disciplines, geospatial big data assumes a pivotal role due to its inherent connection to physical locations and spatial relationships. Several key data sources contribute significantly to this domain. Satellite imagery, acquired from Earth-observing satellites, offers high-resolution snapshots of the Earth's surface and is instrumental in land-use planning, environmental monitoring, and disaster management. GPS data, sourced from vehicles and smartphones, provides continuous streams of location information, aiding in transportation optimization, location-based services, and route planning. Moreover, crowdsourced data collected from citizens through platforms like social media or dedicated apps offer valuable, real-time insights into urban dynamics, public sentiment, and localized events. This type of data empowers citizen engagement and can inform urban planning and decision-making.

The Internet of Things (IoT) has also emerged as a vital source of geospatial big data, contributing real-time sensor information from an array of devices embedded in urban environments. These devices collect data on various aspects, including traffic flow, energy consumption, air quality, and weather conditions. Such sensor data is indispensable for monitoring and managing transportation networks, optimizing energy usage, and responding to environmental changes promptly [10]. In summary, geospatial big data encompasses a unique subset of big data characterized by its continuous updates and inherent geographical attributes. The volume, velocity, and variety of data within this category provide an unparalleled opportunity to gain insights into urban dynamics, infrastructure management, and more. The combination of satellite imagery, GPS data, crowdsourced information, and IoT sensor data offers a comprehensive view of the built environment and its evolution. Harnessing the potential of geospatial big data requires advanced analytical techniques, data quality assurance, and robust data governance. Furthermore, it necessitates the cultivation of data science skillsets tailored to urban challenges, ensuring that built environment professionals can unlock the transformative power of this invaluable resource [11].

Table 1: Key Sources of Geospatial Big Data

| Data Source | Description |
|---|---|
| Satellite/Aerial Imagery | High resolution imagery capturing detailed land use patterns and infrastructure |
| GPS Data | Massive volumes of data from smartphone apps, vehicles, etc. on mobility patterns |
| Social Media/Crowdsourcing | Geotagged posts, photos, check-ins providing localized insights |
| Sensors | IoT sensors measuring traffic flow, pollution, energy use, etc. |
| Smart City Data | Data from transit cards, CCTV, waste sensors, weather stations, etc. |

## Analytical Techniques for Geospatial Big Data

Advanced analytical techniques are required to handle the volume, velocity and variety of geospatial big data. Data mining extracts useful patterns and relationships from large datasets. It encompasses tasks like clustering, classification, regression, and association rule learning. For example, clustering analysis groups different urban land parcels based on attributes like zoning, height, and occupancy [12].

Machine learning applies algorithms that 'learn' from data to make predictions without explicit programming. Common techniques include decision trees, random forests, neural networks and support vector machines. Deep learning is a subset of machine learning based on layered neural networks, enabling analysis of unstructured data like imagery. Geospatial big data is combined with these techniques for spatial predictions.

Other relevant methods include graph analysis to model topological relationships and connections in transportation networks. Agent-based modeling simulates interactions between autonomous agents to analyze dynamic processes. High performance computing leverages distributed storage and parallel processing to handle big data computation. Cloud computing provides scalable, on-demand access to storage and analytics services [13].

## Applications in Urban Planning and Infrastructure Management

Geospatial big data is transforming urban planning and infrastructure management in areas like land use, transportation, resource allocation and citizen engagement. Some key applications are outlined below:

**Optimizing Land Use Planning:** Fine-grained geospatial data of high-resolution satellite/aerial imagery and LIDAR data, along with advanced image processing and machine learning techniques, play a crucial role in enhancing urban land use analytics. Such data aids in informing zoning regulations and creating accurate land use maps. Additionally, the classification of land parcels based on morphology and patterns is made possible through image processing and machine learning algorithms [14]. Furthermore, combining spatial demographic data with mobility data from smartphones and transit systems contributes to improved urban population dynamics and growth forecasts. These integrated techniques support more precise land suitability analysis and demand forecasting, which are essential components of long-term urban planning [15].

**Enhancing Transportation Systems:** Geospatial big data obtained from various sources such as GPS, loop detectors, cameras, and sensors has opened up new avenues for enhancing transportation management. This data, as highlighted by Zheng (2015), presents valuable opportunities for improving the efficiency of urban transportation systems. Real-time tracking of urban mobility patterns, including data from taxis, public transit, and personal devices, plays a crucial role in enabling intelligent traffic management and congestion mitigation, as emphasized by Castro et al. (2013). Machine learning techniques applied to traffic data, as demonstrated by Hong et al. (2016), offer significant improvements in flow predictions and routing strategies. Furthermore, spatial analytics, as detailed by Chen et al. (2016), can identify high-risk areas for accidents and pedestrian-vehicle conflict, aiding in safety enhancements [16]. The combination of smart transit data and citizen inputs, as explored by Müller et al. (2018), facilitates data-driven decision-making for transit planning and operational improvements. These technical advancements in leveraging geospatial big data hold great promise for optimizing urban transportation systems and enhancing overall mobility [17].





**Optimizing Resource Allocation:** Resource planning for urban utilities and services can leverage geospatial data analytics to an extensive degree. High-resolution LIDAR (Light Detection and Ranging) and imagery offer invaluable insights into the urban environment, especially in the context of vegetation mapping and biomass estimation [18]. By harnessing these advanced technologies, urban planners and policymakers can make informed decisions regarding the development and maintenance of urban green spaces. As highlighted by Chen et al. (2017), this approach allows for the optimization of urban greenery planning, leading to a more sustainable and liveable urban environment. The use of LIDAR and imagery in vegetation mapping and biomass estimation is particularly significant because it enables precise data-driven decision-making. Traditional methods of assessing urban greenery often rely on subjective observations, which can lead to suboptimal resource allocation. Geospatial data analytics, on the other hand, provides a comprehensive and objective understanding of the existing vegetation cover. This information is crucial for urban planners as it helps them identify areas with insufficient greenery, facilitating targeted interventions to improve the overall quality of life in urban areas.

Another noteworthy application of geospatial data analytics in resource allocation pertains to waste management. In this context, processing waste bin sensor data plays a pivotal role in identifying efficient waste collection routes. As demonstrated by Khan et al. (2020), the integration of sensor data allows waste management authorities to optimize their collection schedules and routes. This not only reduces operational costs but also minimizes the environmental impact associated with unnecessary fuel consumption and emissions [19].

The optimization of waste collection routes is particularly relevant in densely populated urban areas where efficient logistics are crucial. By analyzing geospatial data collected from waste bin sensors, authorities can adapt their strategies in real-time, responding to changing patterns of waste generation and ensuring that resources are used effectively. This leads to a more sustainable urban environment by reducing the ecological footprint of waste management operations. Furthermore, crowdsourced data, which is becoming increasingly prevalent in the era of smartphones and social media, offers valuable insights into urban dynamics. One such application is the utilization of crowdsourced noise pollution data to enable fine-grained noise mapping and targeted interventions. Guillaume et al. (2016) emphasize the significance of this approach, which allows urban planners to identify noise hotspots and implement measures such as noise barriers and speed limits in specific areas. By doing so, they can address one of the most pressing urban challenges: noise pollution, which has a significant impact on the health and well-being of urban residents [20].

Geospatial data analytics empowers urban authorities to address noise pollution in a more precise and efficient manner. Traditional methods for assessing noise levels often rely on a limited number of fixed monitoring stations, which may not capture the full extent of noise pollution across a city. In contrast, crowdsourced data provides a broader and more up-to-date picture of noise pollution. This data can be used to create detailed noise maps, pinpointing areas where interventions are most needed. By deploying noise barriers or adjusting speed limits strategically, cities can significantly improve the quality of life for their residents.

**Engaging and Empowering Citizens:** Citizen engagement and empowerment play a vital role in shaping the future of urban development and services. In this digital age, citizens are actively contributing to geospatial data through various tools, including crowdsourcing, social media, and smartphone apps. Planners and policymakers can harness this wealth of volunteered geographic information (VGI) to make more inclusive and responsive decisions. Citizen engagement not only improves the overall satisfaction of the population with urban outcomes but also ensures that local concerns are addressed effectively [21].

The incorporation of VGI into decision-making processes is a powerful way to engage citizens and foster a sense of ownership and participation in urban development. As highlighted by Campagna (2014), citizens who actively contribute data or information related to their local environments are more likely to feel invested in the decision-making process. This, in turn, leads to greater satisfaction with the outcomes and a higher level of trust in the governing authorities. One practical application of VGI is the analysis of geotagged social media posts. In today's interconnected world, people frequently share their experiences, opinions, and concerns on social media platforms, often with location-specific data. By analyzing this geotagged content, urban planners can gain valuable insights into the thoughts and preferences of their citizens. For example, patterns of positive or negative sentiment associated with specific areas can help identify areas of concern or success in the urban landscape.

Additionally, reported issues and concerns can be crowd-sourced through dedicated apps or platforms, enabling citizens to raise specific problems they encounter in their daily lives. By collecting and analyzing this data, urban authorities can identify hyperlocal concerns and address them promptly [22]. This approach is particularly effective in addressing minor but persistent issues, such as potholes, malfunctioning streetlights, or vandalism, which can have a significant impact on residents' quality of life.

To further empower citizens, open data dashboards can be created to visually synthesize indicators and provide accessible information about various aspects of urban life. As emphasized by Kitchin et al. (2015), open data dashboards serve as transparent and user-friendly platforms that present relevant data to the public. These dashboards can cover a wide range of topics, from air quality and traffic patterns to public expenditure and crime rates. By making this information readily available and easy to understand, citizens are better equipped to make informed decisions and engage in meaningful discussions about their city's future.

The combination of crowdsourced data with official data sources is a powerful way to enable evidence-based participatory planning. By merging information provided by citizens with data collected by government agencies, urban planners can create a more comprehensive and accurate picture of the urban environment. This approach ensures that decisions are grounded in real-world experiences and needs, leading to more effective and equitable urban development. However, the adoption of VGI and crowdsourced data in decision-making processes is not without its challenges. One of the primary concerns is the quality of such data. VGI and some new data sources often lack established quality assurance processes, leading to issues like incomplete data, inaccuracies, and sampling bias, as noted by Senaratne et al. (2017). To overcome this limitation, urban authorities must implement robust validation and data curation procedures to ensure that the information used in decision-making is reliable and accurate.

Privacy is another critical consideration in the realm of geospatial data and citizen engagement. The collection of detailed mobility data and the use of spatial analytics raise concerns about surveillance and the protection of personal information. As highlighted by Kitchin (2016), it is essential to address these concerns ethically and transparently. Clear data usage policies and privacy protection measures must be in place to reassure citizens that their data will be used for the public good without compromising their individual privacy. Furthermore, the volume and velocity of data generated through citizen engagement, social media, and smartphone apps can overwhelm traditional data storage and processing systems. To effectively manage this influx of data, urban authorities must invest in costly infrastructure for data warehousing, cloud computing, and high-performance geo-analytics, as indicated by Yang et al. (2017). This investment is necessary to ensure that the data can be collected, stored, and processed efficiently and in a timely manner [23].

Additionally, the analytical capabilities required for leveraging geospatial data and VGI in urban planning are a significant challenge. Domains like urban planning demand the development of organizational capabilities for complex spatial data science techniques. However, the shortage of personnel with the required skillsets, as highlighted by Tomlinson (2017), presents a key obstacle. To overcome this limitation, cities must invest in training and development programs to equip their staff with the knowledge and skills needed to harness the full potential of geospatial data and citizen engagement [24].

Lastly, the abundance of data and the myriad analytical possibilities can lead to "paralysis by analysis," where decision-makers are overwhelmed by complexity, as noted by Riggs and Gordon (2018). To prevent this, urban authorities must establish clear frameworks for decision-making that prioritize actionable insights and focus on the most critical issues. Effective data visualization and communication tools can also help streamline the decision-making process. Successfully leveraging geospatial big data and citizen engagement, therefore, requires the development of supportive institutional architectures. These architectures must encompass data governance, storage, analysis, and capacity building of personnel with new digital skillsets. By addressing these challenges and limitations, urban authorities can create a more inclusive, data-driven, and responsive urban planning and development process that benefits both the city and its citizens [25].

## Conclusion

The proliferation of geospatial big data, driven by sources such as smartphones and sensors, has ushered in a new era of possibilities for built environment disciplines. These disciplines, notably urban planning and infrastructure management are increasingly relying on data mining, machine learning, and spatial analytics to leverage hyperlocal insights from vast spatial datasets. The aim is to enhance the efficiency of land use, transportation, resource allocation, and citizen engagement within urban environments. While geospatial big data presents immense potential, it is accompanied by several challenges that necessitate careful consideration. One of the foremost challenges is data quality assurance. Given the diverse sources and formats of geospatial data, ensuring data accuracy and reliability is crucial for meaningful analysis. Privacy concerns also loom large, as the collection and utilization of personal location data must be handled with utmost care to protect individuals' rights and data security [26].

Developing the necessary analytical capabilities is another critical challenge. The sheer volume and complexity of geospatial big data demands sophisticated techniques, including data preprocessing, spatial data mining, and machine learning algorithms. As these techniques evolve, so must the skillsets of professionals involved in urban planning and infrastructure management [27]. The demand for individuals with expertise in data science, particularly tailored to urban challenges, is on the rise. Information overload is an issue that cannot be ignored. The sheer volume of data available can overwhelm decision-makers, making it crucial to develop effective data visualization and summarization techniques to distill actionable insights. Moreover, building the institutional architectures required to support the collection, storage, analysis, and dissemination of geospatial big data is a substantial endeavor. This includes infrastructure, policies, and regulations that facilitate data sharing and integration across various domains and organizations.





In spite of these challenges, geospatial big data is a valuable asset in achieving more nuanced, localized, and real-time insights into urban dynamics. It provides the evidence-based foundation for informed decision-making in the planning and management of urban spaces. By leveraging this resource, cities and regions can develop and optimize their infrastructure, resource allocation, and transportation systems with a level of precision that was previously unimaginable. To fully realize the transformative potential of geospatial big data, it is imperative to make strategic investments in several key areas. First and foremost, data governance must be a priority [28]. Robust data governance frameworks ensure data quality, privacy protection, and compliance with relevant regulations. Furthermore, the development and deployment of advanced analytical tools tailored to the specific needs of built environment disciplines are essential. These tools enable the extraction of meaningful insights from the vast and complex geospatial datasets, facilitating smarter and more efficient urban development. However, the tools alone are not enough. Equally important is the cultivation of a skilled workforce equipped with the new data science skillsets required to navigate the challenges and opportunities presented by geospatial big data. Education and training programs that focus on geospatial data analytics, machine learning, and spatial modeling should be established to bridge the existing skills gap.

## References


[1] J. Eickholt and S. Shrestha, "Teaching big data and cloud computing with a physical cluster," in *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, Seattle Washington USA, 2017.

[2] H. Y. Kim and J.-S. Cho, "Data governance framework for big data implementation with a case of Korea," in *2017 IEEE International Congress on Big Data (BigData Congress)*, Honolulu, HI, USA, 2017.

[3] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Context-aware query performance optimization for big data analytics in healthcare," in *2019 IEEE High Performance Extreme Computing Conference (HPEC-2019)*, 2019, pp. 1–7.

[4] S. Ren, "Improvement of dance teaching method of preschool education major based on big data analysis," in *Proceedings of the 2017 2nd International Conference on Automation, Mechanical Control and Computational Engineering (AMCCE 2017)*, Beijing, China, 2017.

[5] C. Martin-Rios, S. Pougnet, and A. M. Nogareda, “Teaching HRM in contemporary hospitality management: a case study drawing on HR analytics and big data analysis,” *J. Teach. Travel Tour.*, vol. 17, no. 1, pp. 34–54, Jan. 2017.
[6] B. Lepri, J. Staiano, D. Sangokoya, E. Letouzé, and N. Oliver, “The tyranny of data? The bright and dark sides of data-driven decision-making for social good,” in *Studies in Big Data*, Cham: Springer International Publishing, 2017, pp. 3–24.
[7] J. Wu, K. Ota, M. Dong, J. Li, and H. Wang, “Big data analysis-based security situational awareness for smart grid,” *IEEE Trans. Big Data*, vol. 4, no. 3, pp. 408–417, Sep. 2018.
[8] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, “Federated query processing for big data in data science,” in *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 6145–6147.
[9] P. Ameri, N. Schlitter, J. Meyer, and A. Streit, “NoWog: A workload generator for database performance benchmarking,” in *2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*, Auckland, 2016.
[10] A. M. Al-Salim, T. E. H. El-Gorashi, A. Q. Lawey, and J. M. H. Elmirghani, “Greening big data networks: velocity impact,” *IET Optoelectron.*, vol. 12, no. 3, pp. 126–135, Jun. 2018.
[11] S. Mohamed Asaad El Banna and N. Makram Labib, “Using Big Data analytics to develop Marketing Intelligence systems for commercial banks in Egypt,” *MATEC Web Conf.*, vol. 292, p. 01011, 2019.
[12] P. Del Vecchio, G. Secundo, and G. Passiante, “Analyzing Big Data through the lens of customer knowledge management,” *Kybernetes*, vol. 47, no. 7, pp. 1348–1362, Aug. 2018.
[13] D. B. Ventura, “Promoting Sustainability in the Fashion Industry: An Exploratory Study of Fashion Sharing in Colombia,” *ijsa*, vol. 1, no. 7, pp. 1–12, Jul. 2016.
[14] M. Shorfuzzaman, M. S. Hossain, A. Nazir, G. Muhammad, and A. Alamri, “Harnessing the power of big data analytics in the cloud to support learning analytics in mobile learning environment,” *Comput. Human Behav.*, vol. 92, pp. 578–588, Mar. 2019.
[15] V. M. Arora, “Harnessing the power of big data to improve graduate medical education: Big idea or bust?,” *Acad. Med.*, vol. 93, no. 6, pp. 833–834, Jun. 2018.
[16] S. Vakhariya and K. Khanzode, “The role of Big Data in enhancing customer experience in UAE retail,” *Int. J. Bus. Adm.*, vol. 9, no. 6, p. 76, Aug. 2018.
[17] F. Bouchama and M. Kamal, “Enhancing Cyber Threat Detection through Machine Learning-Based Behavioral Modeling of Network Traffic Patterns,” *IJBIBDA*, vol. 4, no. 9, pp. 1–9, Sep. 2021.
[18] C.-L. Chang, M. McAleer, and W.-K. Wong, “Big data, computational science, economics, finance, marketing, management, and psychology: Connections,” *J. Risk Fin. Manag.*, vol. 11, no. 1, p. 15, Mar. 2018.
[19] K. Subramanian, K. P. Joshi, and S. Deshmukh, “Improving forecasting for customer service supply chain using big data analytics,” in *Supply Chain Management Strategies and Risk Assessment in Retail Environments*, IGI Global, 2018, pp. 25–41.
[20] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, “Approximate query processing for big data in heterogeneous databases,” in *2020 IEEE International Conference on Big Data (Big Data)*, 2020, pp. 5765–5767.

[21] "Vol. 4 No. 9 (2021): Int. J. Bus. Intell. Big Data Anal. 2021." [Online]. Available: https://research.tensorgate.org/index.php/IJBIBDA/issue/view/23. [Accessed: 26-Nov-2023].

[22] Z. Bi, Y. Jin, P. Maropoulos, W.-J. Zhang, and L. Wang, "Internet of things (IoT) and big data analytics (BDA) for digital manufacturing (DM)," *Int. J. Prod. Res.*, vol. 61, no. 12, pp. 4004–4021, Jun. 2023.

[23] R. Martoglia, "Invited speech: Data analytics and (interpretable) machine learning for social good," in *2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)*, Haikou, Hainan, China, 2021.

[24] A. Nassar and M. Kamal, "Machine Learning and Big Data Analytics for Cybersecurity Threat Detection: A Holistic Review of Techniques and Case Studies," *Intelligence and Machine Learning …*, 2021.

[25] M. Muniswamaiah, T. Agerwala, and C. Tappert, "Big data in cloud computing review and opportunities," *arXiv preprint arXiv:1912.10821*, 2019.

[26] M. A. Mladin, "Supporting regulatory measures in the context of big data applications for smart grids," *Front. Big Data*, vol. 4, p. 675461, Sep. 2021.

[27] E. Paige *et al.*, "A versatile Big Data health system for Australia: Driving improvements in cardiovascular health," *Heart Lung Circ.*, vol. 30, no. 10, pp. 1467–1476, Oct. 2021.

[28] A. Nassar and M. Kamal, "Ethical Dilemmas in AI-Powered Decision-Making: A Deep Dive into Big Data-Driven Ethical Considerations," *IJRAI*, vol. 11, no. 8, pp. 1–11, 2021.