Sustainable Futures 6 (2023) 100118

Contents lists available at ScienceDirect

Sustainable Futures

journal homepage: www.sciencedirect.com/journal/sustainable-futures

A sustainable smart mobility? Opportunities and challenges from a big data use perspective

Riccardo D'Alberto [a,*], Henri Giudici [b]

[a] Department of Statistical Sciences "P. Fortunati", Alma Mater Studiorum – University of Bologna, Via Delle Belle Arti, 41, Bologna 40126, Italy
[b] Department of Science and Industry systems, University of South-Eastern Norway (USN), Norway

ARTICLE INFO

Keywords:
Smart cities
Sustainability
Big data analytics
Mobility

ABSTRACT

This paper discusses the recent insights on the Big Data role in the sustainability of smart mobility. Systematic Literature Review is applied to scientific publications web repositories retrieving 2,000+ records (years 2010–2022). 83 selected publications are analyzed and discussed in detail considering methods, tools, pros, cons, solved challenges, and pending limitations. The final picture shows significant attention given to Big Data handling/modeling, while yet there is marginal consideration of how such solutions effectively consider the environmental concerns. These, instead, represent the leading priority for improving and ameliorating the smart mobility system sustainably. In this regard, possible research directions are proposed.

1. Introduction

Smart Cities are acknowledged to mitigate the continuously arising issues caused by rapid urbanization [1]. Being increasing worldwide [2], as well as at the European level [3], smart cities' projects are meant to offer their inhabitants the opportunity to improve the quality of life in economic, environmental, and mobility terms [4,5]. Regarding the latter element (according to [6] and in line with the findings of [7]), the perspective of sustainable mobility should be grounded on the development of mobility systems that are as much as possible integrated with the overall sustainable development goals, thus building a *"transport system which meets the society's economics, social and environmental needs"* [8]. Therefore, mobility does not have to solely look for an innovative path, but it also must encompass sustainable development goals. Yet, in the same line, [9,10] stress that only with an interdisciplinary and integrated approach that understands how city inhabitants interact with urban spaces and how the transport system can improve this interaction, it could be possible to (re)route the transport system in a more sustainable path.

The support offered by Information and Communication Technologies (ICT) for such an integrated transformation is well-known [11] pointed out that a smart city is (or should be) a *"multi-faceted meaning"* that connects different characteristics and aspects of people and communities' life with ICT [12] described smart cities as urban areas being i) instrumented (e.g., equipped with real-time data-driven sensors), ii) interconnected (e.g., data is integrated into computing platforms), and iii) intelligent (e.g., it provides complex analytics and modeling tools). These features of the "smart system" engage heterogeneous sources like ICT, the Internet-of-Things (IoT), fixed and mobile sensors, actuators, mobile phones, software, etc. The main, common characteristic of these sources is that they collect, store, and share information and, through data, they foster knowledge on the way that humans interact with the physical world. These sources and the massive use of the interconnected technologies that people employ in their everyday life generate a massive amount of data called Big Data.

The fundamental aspect of Big Data lies in its value creation with information [13]. Indeed, the more insights are extracted from Big Data, the more the human activities and urban patterns are understood which, in turn, makes easier the understanding of the complex smart city systems and their related activities [14]. Due to the high potentialities of Big Data, its role in the smart city context is getting considerable attention from researchers [15–18]. However, when dealing with Big Data in smart cities, a prerequisite is to define the field of application, thus making the analysis more effective and efficient [19]. The main reason is that the way of thinking about Big Data can significantly differ, based on the source of data provision (industries, stakeholders, governmental agencies, etc.). In addition, to understand the role of Big Data in mobility and transportation in the smart city context it is important to familiarize ourselves with the concept of smart mobility which comprises several aspects, ranging from the use of connected

* Corresponding author.
E-mail address: riccardo.dalberto@unibo.it (R. D'Alberto).

https://doi.org/10.1016/j.sftr.2023.100118
Received 5 September 2022; Received in revised form 16 June 2023; Accepted 29 June 2023
Available online 30 June 2023

autonomous vehicles [20], to the more efficient planning and management of the urban (and rural) public transport [21], not disregarding the facilitation of micro-mobility and vulnerable road users, e.g., the urban transit of pedestrians, cyclists, scooters, etc. [22]. Here, Big Data has several interlinked uses, potentially playing a significant role in i) setting up and feeding an intelligent transportation system [23], ii) contributing to planning infrastructural networks and the urban development [17], iii) supporting and controlling the automatic vehicular networks [24], iv) mitigating traffic congestions [25], v) predicting and managing the users' behavior [26].

A large literature exists addressing all these single topics. However, although massive research activity, to overcome the challenges of smart city mobility and its sustainable development goals, it is mandatory to grant attention to the effective role of Big Data and its (co)integrated uses in the smart mobility system. Therefore, it is crucial to understand the impact(s) derived from the use of Big Data on the smart cities' sustainable way of life, here understood as the sustainable way(s) in which (and how much) people move around the city, how these aspects (both) benefit the livability of the cities, to what extent people can contribute mitigating the negative environmental externalities generated by their transfer, and, also, how the latter can eventually provide positive ones. These aspects are underrated. While the application of Big Data to smart mobility often tackles specific practical issues without debating the general frame (where complex challenges are in place), the effective contribution of such an application in terms of sustainability is often merely implicit.

The present paper investigates how Big Data is used in the smart city context to impact urban mobility sustainably. Starting from the research question (RQ) *"Which role has Big Data in fostering the smart city mobility?"*, the Systematic Literature Review method is applied to Google Scholar and Scopus web repositories. The retrieved results (from the years 2010–2022) are then analyzed to explore and discuss the interplay of the different sources, tools, and methods related to Big Data and its use for the sustainable smart mobility. The present work synthetizes what is already known about this topic, specifically focusing on how Big Data is used to organize, manage, and monitor urban mobility, highlighting the main differences (and their implications) among the theories and methodologies applied concerning the topic, not disregarding their usefulness for real-world applications. Although the multifaceted features of smart mobility, the novelty of this work is in the peculiar space given to the discussion on how sustainability is comprised within the investigated studies. Indeed, the results of the literature review show that the relationship among Big Data, urban mobility, and the quantitative assessment of its sustainable impact(s) is rather weak, whether compared with the other aspects of the "smart mobility" topic. We also present research questions to address in the incoming future, discussing the possible directions to undertake to fill the knowledge gaps.

2. Big data

The term Big Data refers to multiple heterogeneous definitions, originally referring to a large amount of information (volume), organized in multiple structures (variety), and collected almost real-time (velocity) [27]. Lately, the *"3V's"* have become *"5V's"* by embracing the features of veracity and value [28] hence, hinting at the profitability that the retrieved information (should) have.

Worldwide, the overall volume of information generated/copied in 2011 was equal to 1.8ZB (i.e., around 10^{21}B), an amount that increased nine times in five years [29], while the volume of data created, copied, and consumed in 2020 reached 59ZB, being expected to reach 149ZB before 2024 [30]. Two main technologies are used for collecting Big Data: IoT [31] and cloud computing [32]. The latter also supports the real-time analysis through parallel processing, memory-based platforms, and offline analysis [33].

Practically speaking, the main pending challenges in Big Data concern data privacy and several technical issues about data acquisition, storage, sharing, and the life cycle of data management and analysis [34]. Theoretically and methodologically, relevant problems lie in data representativeness, heterogeneity, and effectiveness of use [35].

Concerning smart cities and sustainable mobility, the value of Big Data is mainly grounded on the geographical information collected during the data-generating process, as well as on the temporal domain: among the terrific amount of generated Big Data every day, nearly 80% are geo-localized [26]. Therefore, Big Data can help in sustaining the planning of urban areas and infrastructure, ameliorating the efficiency and sustainability of both private and public transport, and monitoring the environmental impacts of mobility.

3. Literature review method and SLR results

For retrieving the information from the state of the art on the subject we applied a hierarchical Systematic Literature Review (SLR) [36]. SLR is based on different sequential steps that include:

(1) Scoping – the identification of the topic of interest –.
(2) Research question definition.
(3) Search planning – deciding about the web sources to be searched, the time range to be scraped, the keywords and the Boolean operators to be used, and whether to include the gray literature or not –.
(4) Identification, i.e., the practical 'running' of the search.
(5) Screening – the data set building –.
(6) Eligibility scraping – the refinement of the data set –.

The latter step is divided into different sub-steps. Namely:

1st) Checking for duplicates in the data set; discarding the non-entirely-English-written research works (e.g., those with English abstract but non-English text) and the unpublished ones.
2nd) Trimming the results, i.e., filtering based on subject areas.
3rd) Research works that cannot be found, nor just in-preview mode, are discarded.

By following each step, we defined the RQ: *"Which role has Big Data in fostering the smart city mobility?"*. We identified two worldwide repositories: Google Scholar and Scopus, where we searched the time range 2010–2022. The search was based on the following keys:

- big data, smart city, mobility;
- big data "AND" smart city "AND" mobility.

The gray literature was not considered. The results were exported in a .csv file and scraped for eligibility. As per the Scopus indexing, we discarded the publications categorized in "Economics, Econometric and Finance", "Health", and "Medicine".

Supplementary material (available upon request) depicts all the publications retrieved (with the related information). Fig. 1 depicts the SLR process adopted here, as well as the results in terms of retrieved records at each step. Fig. 2 shows the per year distribution of the 2022 publications retrieved at the end of step 5 and after the completion of the 3rd sub-step of step 6 (n=83, which are the publications considered here for the analysis and hence, for discussing the literature review results).

Among the final 83 publications, scientific articles are 77%, conference papers are 15.66%, book chapters or books are 6%, plus "other" publication types.[1] Fig. 3 depicts the word cloud of the keywords (up to 5 keywords) adopted by the publications.

[1] The publication's editor is IEEE for 43.37% of the publications, while Elsevier and Springer represent 16.87% each; Taylor and Francis are 6% (all the others are very heterogeneous).

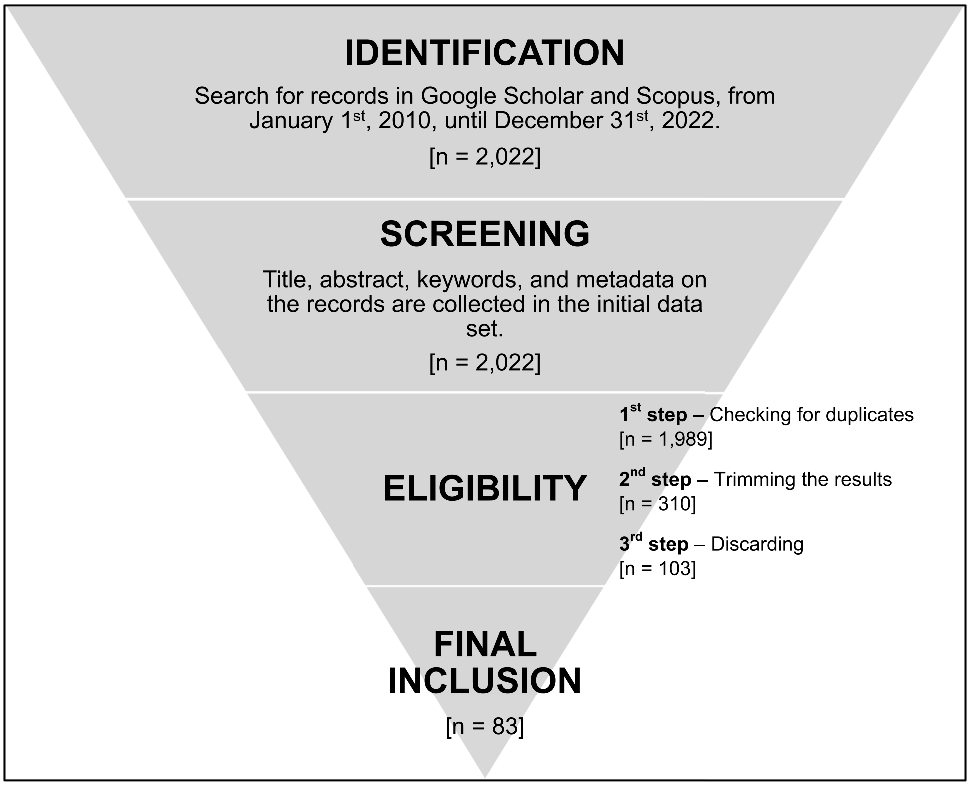

Fig. 1. Systematic Literature Review procedure and number of retrieved publications.

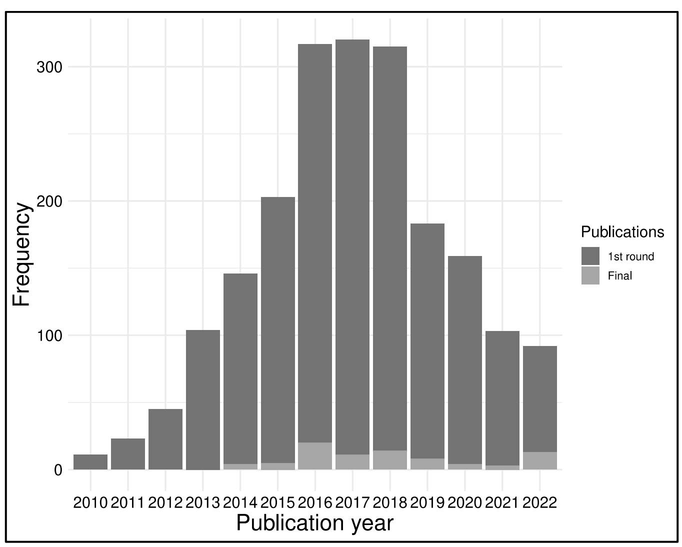

Fig. 2. Per year distribution of the publications retrieved.

4. Sustainable framework from (SMART) city and (SMART) mobility

This section presents the results of the performed SLR, framing the results as shown in Fig. 4 and Fig. 5. In this study, the understanding of how to improve the sustainability of smart cities needs to undergo the understanding of urban areas and smart mobility. In these aspects, the use of Information and Communication Technologies, data analytics as a service, software as a service, and cloud computing play an important role in the analysis of (Spatial) Big Data. In what follows, we discuss the works that focus on how Big Data is used in smart mobility and urban areas, thus fostering a reflection on how the sustainability concept can be integrated into smart mobility one.

4.1. Urban areas: design and planning

The main priority of a healthy (smart) city is the provision of high livable standard levels for its inhabitants which, in turn, (it is supposed to) make the (smart) city more sustainable *per se*. In this context, (Spatial) Big Data can significantly play a major role in the sustainability of smart mobility [34] describes and discusses the potentialities and challenges of Big Data and related analytics (i.e., the collection of dedicated software tools and database systems run by machines that have a huge processing power). The author also highlights the

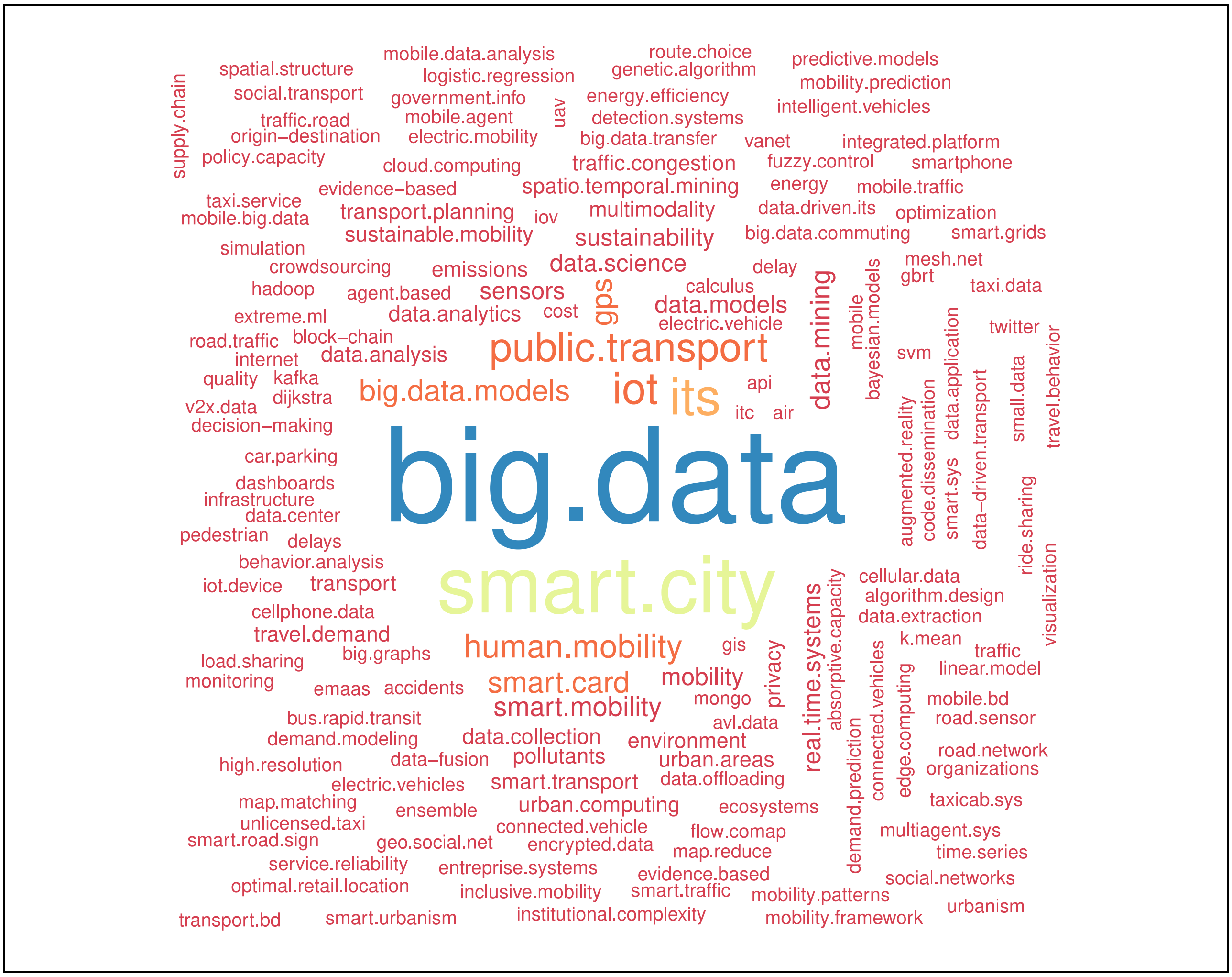

Fig. 3. Word cloud of the publications' keywords.

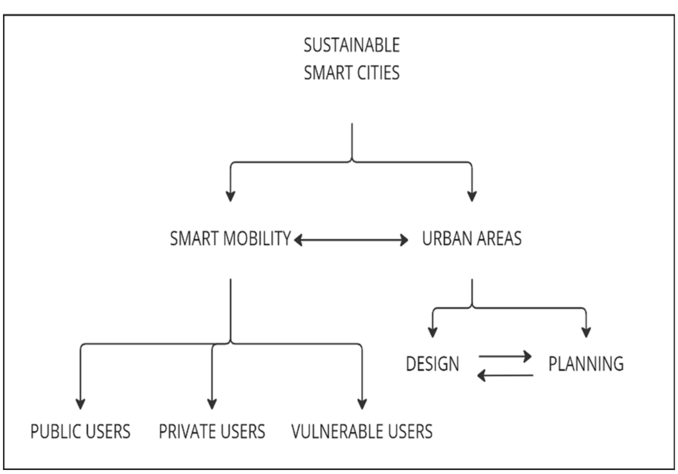

Fig. 4. Diagram framework of results.

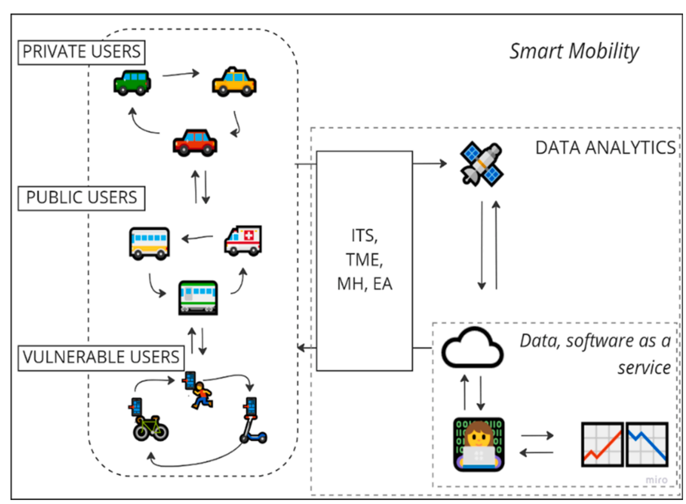

Fig. 5. Smart Mobility Architecture model from results.

increasing role played by Big Data in changing the design and in planning, managing, and operating the urban environment, stressing that such potentialities are not limited to one research field, but do cross various disciplines. Hence, the discussion extends to the development of (smart) sustainable urbanism and, also, to (smart) sustainable mobility. Nevertheless, the skills and knowledge of the proposed solutions are yet extremely fragmented. In addition, there is an overall weak connection between cities' smartness and cities' sustainability. This weakness mirrors the soft connection(s) among different disciplinary approaches to the common challenges offered by the analysis of the problems of smart cities, increasingly tackled just through a data-driven approach. The core suggestion of [34] is that future, sustainable urbanism projects and initiatives should be addressed through stronger interdisciplinarity: only this vision could lead to dealing with the issues and challenges of the smart cities' everyday life. For example, only an efficient smart mobility system that fosters sustainability by prioritizing, e.g., walkability, cycling, public transportation, etc. will ameliorate the urban environment. In contrast, the effectiveness of smart mobility can be factually fostered only through a conjoint approach to transportation, environmental, and land-use planning which, in turn, would not be possible by disregarding the analysis of the individual user's behavior.

In this sense, [37] offers relevant insights by demonstrating how much system dynamics integrated with Big Data platforms can help to improve urban sustainable mobility. Their conceptual model is based on system dynamics and casual relationships among all the factors related to mobility users' needs, putting the latter at the core of the analysis. This approach is particularly effective in addressing specific demand segments within the whole mobility system, pointing at the possibility to meet the challenges of an intermodal solution using smart technologies. These can be embedded in all the phases of people's journeys, from planning to execution, whereas they can generate data that further nurture the model. In a sort of loop, such a decision-making tool can be very useful for policymakers, private companies, and final users, maximizing the accessibility of information and access to transport means, minimizing consumption, mitigating socio-environmental challenges, and improving life quality. Further developments should involve the adoption of a dynamic approach based on real-time databases, the openness of interfaces, and the promotion of data transferability. In this regard, for example, strict policy measures could be adopted in smart cities where the application area involves critical societal infrastructure.

The central role played by people's behavior with respect to urban mobility emerges also in the literature review by [38]. First, by presenting the latest advances in transport modeling and by describing the recent transport model approaches, the authors find out that activity-based models and agent-based models are particularly accurate in describing human mobility, mainly because travel demand can be always maintained at the level of individuals/individual action(s) throughout the model. Second, they discuss how to extrapolate information about mobility behavior from big databases on trips and the reason(s) for the planned activities. Whereas future research hints at modeling the potential interactions of the different transport modes and mobility users, an additional effort must be done to develop robust algorithms that can build coherent individual schedules from sparse mobility traces. In this direction, the adaptation of the existing approaches to the network-based methodology (e.g., Relational Markov Networks, Hidden Markov Models, etc.) is strongly envisaged.

From the practical point of view, the limitations envisaged by [34] were explicit in [39] who discussed how the sustainability of smart cities is constrained, on one hand, by the low level of interconnectedness existing between the different departments of the city municipalities and, on the other hand, by the lack of compatibility among the different data that are collected. In this sense, the key role of Big Data in advising transport researchers is supposed to go through the fundamental re-assessment of which kind of transportation-valuable information is extracted, and *extractable*, from such kind of data [40] proposes to use robust predictive models and real-time planning to overcome the challenge but, conditional to the deep and detailed re-thinking (and re-structuring) of the whole *"Big Data paradigm"*. The author points out the downside of each potential opportunity that Big Data conveys. For example, there are no guarantees that the same data will continue to be available in the same form forever, hinting at comparability issues. In addition, there are well-known privacy issues behind the inherently invasive nature of big databases, as well as the less well-known ethical issues behind the control of such new data. Moreover, there are potential energy implications related to the generation, storage, and transmission of Big Data (e.g., what if these implications are comparable to the energy problems deriving from physical movements?). Skepticism concerning full data-driven solutions is shown by [39] e.g., about the effort made for using Big Data in urban policy-making for carbon emissions mitigation. Considering this issue, the author demonstrates how the answer to the problem lies in the use of *effective* technological solutions rather than in the intrinsic value of data collected/available.

The survey presented by [25] explores the paradigm of Big Data and how Big Data and its life cycle are tackled by research and industry. In their work, the authors present open challenges and future research directions in which they stress the importance of, among others, aspects such as real-time data analytics, security and privacy for transport users, data sources from new technologies, futuristic transportation solutions, and new intelligent transportation systems.

4.2. *Smart mobility*

This section describes the components of smart mobility as shown in Fig. 5, facilitating the framing and the discussion of the SLR results. It is not surprising that a considerable number of research works target (Spatial) Big Data, IoT, cloud computing, and software analytics, while an optimistic approach hints at how such innovative tools might bring mobility into a sustainable path. Framing the advancements proposed in the last decade with data analytics as a service and software as a service, this section discusses the addressed challenges, the unsolved ones, and the pending limitations of such works.

4.2.1. *Intelligent transportation systems (ITS)*

[41] provided a detailed description of i) mobility data, ii) the tools that contribute generating them (e.g., smart cards, mobile phones, GPS, sensors-based devices, etc.), and iii) the existing methods and procedures for their management and analysis (e.g., data clustering, data mining, pattern mining, location-oriented queries, trajectory-oriented queries, etc.). According to the authors, there are still some issues related to the practical deployment of Big Data, for example, the creation of moving-object databases and the management of semantic-aware data (with the related privacy claims constraints). The latter is presented as the priority to be addressed and solved for the achievement of the mobility system of smart cities.

Comprehensive literature reviews on the role of Big Data in the Intelligent Transport System (ITS) are carried out by [23,42,43]. These authors described the source(s) of data, the ways of data collection and data analytics, the functioning of different tools and integrated platforms, and their specific uses. Moreover, they presented and discussed several case studies where the combination of Big Data and ITSs contribute to ameliorate the efficiency of the transportation sector [23] propose a historical review of the relations between Big Data and ITS, stressing that data inaccuracy/incompleteness, privacy, volume, and processing timeliness are still current and urgent issues to address in the incoming future. However, the authors highlight another relevant problem that should gain attention as soon as possible: the openness of data. This impelling priority represents much more than the solely *"open data idea"*: it is envisaged that transportation sector users effectively gain access, and interact, with the information provided [42] propose a conceptual shift in the concept and role of Big Data by suggesting that Big Data should be perceived more as 'technology' itself, rather than a mere 'product' generated by different technologies. In contrast, [43] focuses on the adoption of Computational Intelligence (CI) models, shifting the paradigm of thinking and using Big Data for evolving ITS radically. The author discusses several modeling challenges related to CI and optimization, e.g., the gradual shift from models to analytics; the shift from *"data batch to be processed"* to continued analysis in real-time; the role of data representativeness, and the implications of correlations that will never be 'enough' compared to causation. The idea conveyed is that the future of ITS will be built on the unique identity of humans and machines, mutually supported by systematic statistical thinking. As a direct consequence of such a revolution, only the interdisciplinary approach will enhance the explanatory power of CI applications.

Big Data handling and mining, the structure of Big Data sources, the mobility models, and their relative use in terms of ITS features and existing analytics approaches are topics discussed in general terms by [44,45]. The former work provides a conceptual map that classifies the pending challenges related to the different phases of Big Data use in smart mobility, e.g., data preprocessing and cleaning, modeling choice, and validation of predictions. From the methodological point of view, the focus is on deep learning techniques, while in terms of the information to be used, the focus is mainly on GPS data. Overall, the elected approach is that of data mining. This, indeed, could boost the possibilities offered by the combination of transport data and social transportation data, by envisioning a physical and cyberspace finally conquered by the information produced by the social signals.

The use of multiple integrated sources for facilitating real timeliness and efficiency of data mining procedures aiming at ameliorating traffic data analysis is proposed by [46]. By reviewing the literature on traffic data applications, the authors test the potentialities of the A*STAR data analytic platform that is developed for integrating, storing, and managing data sets from multiple sources. In a complex ITS environment, like the one in Singapore, this platform represents a successful case study in terms of data merging, storing, access, and sharing. Similarly, [47] present the City Administration Dashboard, a data analytic platform designed for a wide variety of users (from basic to highly skilled) built for public transport analysis and planning. The application was developed, and tested, to satisfy the traffic requirements of the city of Curitiba (Brazil). Namely, this application was developed to adapt dynamically to other urban contexts, by enhancing three main benefits, such as flexibility, generality, and replicability (which in turn reflect easy adaptability to other cities).

4.2.2. Traffic monitoring and eco-routing (TME)

[48] explored the use of low-cost IoT solutions that can be regularly displaced over the road, thus detecting vehicles. The main goal of this work is to improve traffic management systems by acquiring data, streaming it in real-time and analyzing it. The final output that should be obtained is a user-friendly interface that analyzes the traffic with relatively dense characteristics, also providing alternative solutions in case of jams (hence, easing traffic congestions and, at the same time, being eco-friendly). In a broader context, [49] developed a hybrid traffic simulation model by combining pre-set and real-time data at multiple intersections. Equipped with smart detector devices, the information collected allows us estimating different traffic volumes, signals, trajectories, and the relative emissions produced. The result of the work is an architecture controlling signals and vehicle volumes, generating vehicle trajectories, and predicting travel times but also emissions performance. Similarly, but with a more specific methodological focus on genetic algorithms and fuzzy systems, [50] discuss how the use of an ITS architecture equipped with RFID sensors can be adopted for the automatic control of traffic lights at intersections. By combining genetic algorithms with fuzzy control methods, the authors were able to control and reduce the waiting time at the intersection from traffic vehicles. This work envisaged as future research directions three aspects: the real-time readings from the adopted sensors, the scheduling time by mathematical approaches, and the efficiency of the related traffic conditions. Yet in the same line, [51] elaborate a prototype architecture for real-time traffic control, the Simulation of Urban Mobility (SUMO) platform, that is integrated with the Kafka tool and Spark pre-processor which allow using large bandwidth of incoming (real-world) data that can be promptly analyzed but also integrated with multiple data sources. SUMO proved its flexibility and efficacy; in addition, it does not require data quality and plausibility checks. The challenge of performing road data analysis in real-time (also, by including sensitive data) is tackled by [52] by developing a model for improving the efficiency of smart analytics transportation. This mitigates traffic issues, and, in addition, it solves the well-known privacy concerns. By combining Spark and Hadoop tools thus, by using both platforms, a layered architecture system was developed, and it was validated with real data at high accuracy. Hadoop is an open-source framework that processes very large amounts of data for developing and executing various distributed applications and it is used also by [53] for investigating smart metropolitan mobility. Graphs and indicators describe traffic conditions in real-time by exploiting the information retrieved by ITSs and IoT such as road sensors and vehicular networks. This study evaluated the potentialities of real-time algorithms while making decisions leading to the conclusion that such an approach can be adopted for any transport problem. The real-time algorithm was validated using online traffic data from different countries such as Germany, Spain, and Denmark [54]. Focus on traffic-related indicators by combining data on vehicle speed, traffic flows, and accident information. They use these heterogeneous indicators for predicting urban traffic in real-time, by resorting to use Big Data by adopting deep learning techniques. The latter allows us to automatically discover the most representative features affecting traffic status from a huge amount of data, without roughly approximating the unknown relations among all the impacting factors (as it was, instead, in the case of prior statistical and machine learning methods). However, future research must focus on solving the problem of the joint optimization of multi-source data and traffic modeling as well as the issue of conventional computing that is inadequate in considering hundreds to thousands of layers (and thus parameters) needed for prediction. In the context of traffic congestion improvement through an ameliorated Vehicle2Vehicle and/or Vehicle2Anything communication system, [55] investigate the displacement-aware Big Data endowment to improve the accessibility of requested data from vehicles when there is occurrence of high traffic density and/or in case of remote driving locations.

In a broader context, [56] analyzes the combined use of urban intelligent transportation tools and technologies such as GPS and GIS, which can be applied to solve traffic congestion problems. In this work, the author adopted the mentioned data and collected/provided information sources which were integrated with cloud computing and clustering techniques to retrieve a reasonable amount of coherent and punctual information that, in turn, is important to provide the road user with diversified services at high reliability [57] adopted a Vehicle2Vehicle architecture based on IoT, but also a Vehicle2Infrastructure communication solution with Big Data retrieved in real-time. The multi-faceted system (named Cassandra) is validated, showing a very positive response in leading drivers to understand the traffic congestions and consequently adopt effective solutions that, in turn, lead to shorter travel distances and lower CO_2 emissions. Also, [58] studied a scalable real-time system capable to understand traffic flows, travel behavior, and spotlight traffic occurrences, accidents congestions, etc., deploying both road sensors and GPS sensors. The presented system was validated with a case study from Casablanca (Morocco) and large data sets. The main finding obtained was an optimized traffic time, basically resulting from the lower time spent by vehicles in traffic. Traffic congestions are also tackled by [59], who discuss how improved mobility solutions can benefit the lifestyle of Indian people as well as the relative states' roads. The investigation of [59] prioritized the discovery of short distance travels by using traffic patterns and travel durations combined with the Support Vector Machine (SVM) and regression analysis. To understand the validity of the approach carried out, this work used traffic Big Data for the prediction of vehicular congestion conditions. Yet in the same line, [60] combined Big Data and machine learning techniques with Augmented Reality (AR) to predict traffic conditions. In their study, the assessment of information reliability and the whole analysis are performed to detect and understand specific traffic events.

[61] propose solutions based on Google platform by discussing strengths, weaknesses, and commonalities of the existing framework for modeling and managing traffic, not disregarding the key objective of travelling time reduction. The most addressed path for settling these challenges is the mapping of the traffic patterns by means of the extraction and modeling of mobile phone data. This is particularly viable, above all, in metropolitan (dense) areas. By integrating the best

existing world solutions, the authors succeed in ameliorating traffic safety and smoothness. In the context of improving the traffic congestion in the city of Riyadh (Saudi Arabia), [62] present a traffic management system that adopts an IoT sensor for counting road vehicles and uses cloud computing to store in real-time traffic information which is shared at the same time with the road users aware of the upcoming traffic conditions. The proposed architecture provides real-time information regarding traffic intensity [63] proposed to predict large-scale cellular networks and mobile traffic, using predictions to feed model-based traffic planning frameworks, also by resorting to time series approaches. Data obtained from 9000 cellular towers were adopted in this study extracting and modeling traffic patterns using time series predictions to assess future traffic movements. Similarly [64], investigated the pathways to the collection of cellular network data, by specifically focusing on the problems of data dissonance and loss. A novel, improved algorithm was proposed; it leverages smart cards transactional records and vehicle location records to expressly verify the validity of data, finally benefiting IoT and its influence on public transportation in smart cities. Analogously [65], present how the use of smart cards can be adopted to improve the mobility services of a city, considering the municipality of Sydney (Australia). Indeed, the latter set its ambition to make important municipal services (e.g., hospitals) accessible within 30 min to every citizen in the urban area. Among the challenges envisaged from the authors, the main, relevant one is linked to the accessibility of Big Data repositories. In fact, these are equipped by commercial companies, thus ending up in being not fully accessible.

4.2.3. Mobility of human (MH)

[66] discussed the relevant shift that has occurred in the last years in the analysis of human mobility. At first, the latter has been approached by transportation researchers but, recently, it has become a broader interdisciplinary research topic, being of interest for computer scientists up to physicists. The effect of the diverse research approaches lies in the involvement of different theoretical grounds, types of data, and methods that have been recently applied to study and investigate the same (or strictly close) research questions. Consequently, the authors suggest integrating several research views for stimulating a cross-interdisciplinary discussion on the mobility of humans, not leaving aside the behavioral features as well as the different factors influencing it [67] extrapolate human mobility data from Twitter, using comprehensive process mapping for studying the correlations between human activities and their mobility. An application to the case study of New York (USA) shows that the use of such social media produces results aligned with the existing human mobility research, being a valuable and viable resource for studying human mobility at the city scale. Twitter data integrated with information retrieved by metropolitan smart cards for analyzing the passenger traffic flows, e.g., in special cases that are linked to unusual meteorological phenomena, accidents, public gatherings, etc. are particularly useful in offering insights on passengers' flows in real-time. In addition, to define and make accessible the information on public networks, three visualization platforms are developed: HeatMap, AnimatedRibbon, and TweetBubble views.

Strictly connected to the mapping practice, geolocation is emerging as a key aspect of human mobility analysis. 'Geopositioning' prediction embraces at least four components of challenge, i.e., i) the collection of data from different sources (e.g., GPS, Wi-Fi, etc.); ii) the detection of popular geolocated areas by means of data-density and/or time clustering or personal trajectories mining (that is also based on time); iii) the geolocation with event-based data; iv) the prediction power of geolocalization models [68,69]. Among the statistical-based models, the most celebrated ones include (but are not limited to) Markov-based approaches, Bayesian network/regression-based methods, and neural networks-based solutions [70] presented an innovative proposal for investigating people's mobility derived from the detection of popular geolocations within the city area. Indeed, they resort to retail store placement information and, by collecting and processing the whole set of information related to popular city areas, they infer the optimal area where a potential store attracts more customers. Similarly, [71] propose a framework for mobility analysis with Big Data that is based on two steps. Frist, there is the humans' mobility mapping. Second, the popular spots in the city are identified and selected. By exploiting the geo-positioning of individuals in, e.g., social media, users' trajectories are then reconstructed.

[72] elaborated a tool (that is also a framework), named Allboard, which uses mobile phone data to perform transit and optimization analysis, particularly useful in assisting the mobility design and planning operated by direct users. Nevertheless, it can also help decision-makers, while confronting different scenarios. Strengthening the data-driven approach with real-time analysis solutions, [73] use mobile phone Big Data to develop a transit network design that, first, does map the travel journeys by using frequent anonymized patterns from mobile user operators. Secondly, optimal travel routes with relative frequencies are considered and offered to the final user. Moreover, by combining both these steps in real-time, the method is suited for presenting the routes which optimize the travel users' resources (*in primis* time, but also energy). Here it is understood how human mobility, through mobile phones, can be assessed. However, the arising use of IoT can also provide useful information in terms of mobility [74] studied how IoT and human mobility patterns differ by adopting an extensive data analysis. They collected data at the national level in China, where mobility information from 1.5 million IoT and approximately 0.4 million phones were retrieved and analyzed. Spatial and temporal analysis from the retrieved data sources were performed finding that both sources differ. In addition, they predicted IoT information and, also, investigated the prediction limits.

4.2.4. Environmental aspects (EA)

[75] Presented a comprehensive literature review highlighting the Big Data role in multiple 'green concepts' such as environmental sustainability (for example, with respect to carbon emissions, pollution, and climate change), energy management in industrial facilities, buildings, and infrastructures, overall sustainability in terms of renewable sources use and nature protection/conservation. In these regards, [76] elaborated a methodological solution to predict in real-time the polluted areas, based on the integrated use of Big Data and k-mean clustering methods. The authors stress that the ability to predict and detect densely polluted urban areas and, for example, in differentiating them from the cleanest ones will have direct consequences on the sustainable management of the environment. At the same time, relevant consequences are linked to the assessment of the city's livability in terms of health (and wealth) or, in other words, to the extent that the city can be inhabited. Similarly, [77] demonstrate the beneficial use of Big Data in ITS for monitoring vehicular emissions, thus improving the management of cities' air quality. The combined use of ITS and traffic data can provide indeed high-resolution information on traffic dynamics 24 h/day in the case study considered (Nanjing, China). The detection of emission patterns can serve, in turn, for investigating traffic restriction solutions under different traffic scenario(s). Then, for example, by combining the management system with weather and atmospheric models, the former could help in mitigating the impacts of air quality, ameliorating its management.

4.2.5. Public transport

[78] Proposed to optimize the supply of public transport service using the Automatic Vehicle Location (AVL) data in the Netherlands. AVL provides information on the arrivals and departures (all along the journeys) of the public buses. By integrating these data in the GOVI-tool and by analyzing them in real-time, the amelioration of the service quality as well as the reduction of the negative effects of traffic bottlenecks is accomplished. It is worth to mention that only already available data are exploited and, hence, new data collections are avoided. The GOVI-tool is developed *ad hoc* thus, it explicitly considers this issue: it is

fed by retrieved data from all the public transport operators [79] moved forward the framework by integrating urban transportation data with mobile phone data. Such a strategy allows developing a model that estimates the travel demand over a time window of one hour. The results hint at offering an almost real-time mobility frame that has the main pro of holding even with large masses. The usability of the transportation system is ameliorated, and, in turn, this enhances the users' travel demand. Similarly, public transit data are used in combination with geolocalization data by [80] who enhanced the study of people trajectories within the public transportation system by estimating the first-and-last mile traveled pattern. In Brisbane (Australia), [81] tested a new geovisualization-based method that exploits smart card data and helps examine the spatial-temporal behavior of busses users within the Bus Rapid Transit (BRT) framework. The model proposed investigates users' temporal dynamics by comparing the BRT trajectories with other busses trips on working days and holidays. The results show that BRT efficiency is improved by means of evidence-based planning which, in turn, increases the general efficiency of the public service.

A management system that improves the quality of public transport despite the limited financial resources that cities administrations have at their disposal is proposed by [82]. Limitedly to the urban busses transportation, the system is based on three operational steps: first, data from busses trips and trajectories are collected; second, analytical approaches are implemented to detect users' patterns and busses delays; third, an interactive output visualization is elaborated and offered to support decision-making as well as the individual decisions (the system was validated in Fortaleza, Brazil).

In line with these approaches, [83] exploited internet data to feed a prediction model for the assessment of public transport arrivals at social events, e.g., concerts and sports games at stadiums and arenas. An off-the-shelf technique proved to practitioners that the model is efficient and robust in predicting trip timing in the successful case study of Singapore. The case of Lisbon (Portugal) is presented by [84] analyzing a context-aware multimodal traffic analysis oriented to public transportation and the latter's relative efficiency. After presenting their data-driven decisions, the multimodal (heterogeneous) data-driven approach is presented and applied to the case of Lisbon with a spatio-temporal analysis. The work concludes with the advice from the authors to adopt the presented model to improve the inclusivity and transparency of the transport authorities while performing decision-making. Yet in Lisbon, [85] adopted a similar spatiotemporal approach to propose a novel real-time multimodal pattern approach for the improvement of public transportation (bus and metro) including walking distances and waiting times from the users. This is done by integrating inside the dynamic Origin-Destination matrix some statistical information, and transport users' information, finally proving that nearly 77% of the transportation can be predicted at reasonable levels of accuracy.

4.2.6. *Private mobility*

The fundamental role of taxicabs as a door-to-door service is described by [86], who also put them in relation to the smart mobility system. By envisaging a profound shift in the idea of taxicabs as private road users and promoting their identification as (potential) pieces of the public transportation network, the authors stress that taxicabs can be seen at the same time as a vehicle and as a system. Their peculiar semi-private offer can allow the bottom-up rethinking of a system in which such vehicles can adapt to both users' demand and urban features. This should progressively lead them to become an integrated component of the mobility system of the city, thus positively impacting its sustainability. In this sense, [87] compare the information from the smart card transactions and the trajectories made by the geopositioning of taxicabs, exploiting these data to assess travel patterns. The comparison is made possible by an analytical framework that was designed and applied to a real case study (Singapore). Results hint at the high correlation existing in the spatial distributions of the travel demand extrapolated from the two transport solutions but they also help in detecting relevant differences in terms of travel distances and spatial interactions of the communities liked to transport modes at different times [88] elaborated OnATrade, a method for the real-time detection of the anomaly trajectories of taxicabs, aiming at exploiting GPS information for planning real-time the optimal travel trajectory (at first) while detecting the anomalies during the driving (secondly). At first, a route recommendation algorithm selects candidate routes for the journey and, once it selected the route aimed to drive, at the second step it detects the ideal position of the taxi in contrast with its actual position spotlighting possible anomalies in the journey. From the point of view of taxicabs passengers, [89] investigated the users' demand by developing the so-called Dmodel which uptakes the taxi as a whole system of sensors. These are embedded in the model itself, which is then fed by a very large amount of data. The resulting complexity is mitigated by a novel parameterization approach based on the entropy of pickup events. Evaluations of the model have been made on real-world data proving that the proposal outperforms traditional approaches, being also potentially further enhanced by customized online training. A different scope for the use of Big Data in ITS was proposed by [90] who investigate the identification of unlicensed taxicabs with a model which comprehends two components: 1) the candidate selection and, 2) the candidate refined. This model allows screening the candidate unlicensed taxi(s) that is (are) then sent as input for the second component which, based on machine learning techniques, finally refines the candidate list. The model validation is based on real-world data from China, proving to be versatile in unlicensed taxis detection with data-driven ITS, even if the model learning phase and accuracy could be further enhanced.

A peculiar focus on the technologies and methodologies that can convey Big Data in relation to private road users and vehicles, characterizes the research of [91,92]. They focus, respectively, on Big Data in Vehicular Ad-Hoc NETwork (VANET) and VANETs in Hadoop. The former work proposes to use machine learning methods to detect the misfunction condition of VANETs, efficiently ameliorating their performances. The latter adopts the Dijikstra algorithm to track down routes for VANET in Hadoop, demonstrating that processing time can be greatly reduced by using certain precautions (e.g., by augmenting the cluster nodes in the framework). A similar approach is adopted by [93] studying the relationships between Big Data and the Internet of Vehicles (IoV). They investigate how massive data are stored and transmitted by IoV, assessing the beneficial role of Big Data in IoV by exploring the potentialities of the autonomous vehicles (AV) but also the whole system's capacity in detecting or predicting traffic flows [94] deepen the potentialities of using network calculus in querying the network models. In contrast with conventional optimization approaches, the carried-out methodology adopts a multi-objective optimization based on traffic parameters such as traffic waiting time, travel time, and supply of services for AV. Results hint at higher satisfaction for the autonomous vehicle service, with rather accurate estimates provided to customers in relation to AVs service waiting time. An issue related to AV networks and ITS is represented by the huge amount of data gathered by the sensors which the vehicle is equipped with. Whereas these sensors can potentially collect tons of data, the latter should be stored; their massiveness constitutes a relevant limitation to their potential use [95] describe a data-transfer framework that, based on volunteer vehicles, allows to transfer data from several data centers through the vehicles themselves. The proposal is extensively validated in real-world scenarios showing significant results in terms of framework efficiency, vehicle capacity of transporting the information that they directly collect, energy saving, etc. [96] approach the same context, i.e., the integration of electric vehicles in electric-Mobility-as-a-Service (eMaaS) systems, but with a different perspective. Specifically, by studying how to improve the interoperability of different data sources with respect to mobility in smart cities based on the combined use of Enterprise Architecture (EA) and Application Programming Interphase (API). EA and API enhanced the creation of an architecture focused on the interoperability of data within the eMaaS. The proposed architecture comprehends at its core, steps such as i) context understanding, ii) service layers, iii) business layer, iv) application and data processing layer, v) data space layer, vi) technology layer, and vii) physical infrastructure layer. According to the authors, the proposed architecture provides a useful tool for sharing and deploying the mobility data of electric vehicles in smart cities [97] presents her perspective with regard to AV-algorithms, mobility technologies, and transport design/planning. Through a statistical analysis of different mobility data sources, the author concludes that the behavior of human-driven vehicles might act differently from what is expected from AV-algorithms although the AV can improve traffic operations, for example, by improving the maneuverability and route trajectories of AVs.

In the context of ride-sharing networks, [98] proposed a suitable solution to a unique data privacy concern. Indeed, if the organization of ride-sharing is supposed to have several pros that benefit the smart city mobility and the environment, sensible information about pick-up/drop-off locations as well as trip time and routes are supposed to be shared, thus undermining the sustainability of information exchange. Based on a similarity measurement technique that pairs the users who have close characteristics (which, in advance, have been encrypted in binary vectors), the authors demonstrate that ride-sharing can be coordinated preserving the privacy of the user, but also guaranteeing short search time and higher precision compared to other approaches.

Considering the facilities connected to private road users, [99] investigated the availability of smart parking structures in smart cities. They propose a Gradient Boostil Regression Trees model (GBRT), based on multi-source data, whose algorithm predicts at regular intervals the car parking availability (both for regular days and in case of special events). The results show that, by using multi-source data and the GBRT method, the prediction error is largely reduced.

Peculiar mention should be destinated to the Unmanned Aerial Vehicles (UAV), as per the ones inspected by [100]. Being still uncommon vehicles and, for sure, not properly definable as *"private mobility users"*, such vehicles proved to be a low-cost but effective driver of code spreading over the road network, merely acting as 'mules'. Indeed, the method proposed by the authors is low-cost but quick in terms of network coverage, representing an effective solution to the very long trailing issue.

In addition, also electric vehicles (which are increasingly seizing smart cities worldwide) play a specific and relevant role that requires a special mention among private road users [101] describe the role that they have in smart cities and the potentialities of data that are generated by them. This work reviews existing data analytical tools needed to integrate electric vehicles into the electric grids. After presenting how electric vehicles can be connected to the electric grids, the authors describe the related challenges, requirement analysis, and challenges needed to be addressed. Straightforwardly, these vehicles contribute to increase the sustainability of smart cities with the most evident impact resulting from carbon emissions lowering. In the same context, [102] propose a multi-layer architecture that can store, process, and analyze in real-time Big Data collected and generated by electric vehicles. Such architecture improves the interoperability between the different infrastructures that are needed for enhancing eMaaSs. In addition to offering guidance for municipalities and policymakers in improving eMaaS, the proposed architecture can support transport companies in deploying the electric-Mobility-as-a-Service in smart cities' areas, also generating data that are openly available for end users. In this regard, these data help create value-added services to improve citizens' quality of life. About the topic of multi-layer architectures and, precisely, with regard to multi-agent systems, [103] adopt such architecture to improve the communication between road infrastructure and autonomous vehicles. The correct usage of the mentioned architecture can lead to a dynamic monitoring of road traffic. In this work, the architecture as well as its functionalities are explained. A narrower focus on the electric vehicle owners and users is offered by [104] who propose a Mobile Edge Computing (MEC) framework for assisting them with data-driven solutions that are directly communicated from vehicles and/or to vehicles. A trivial example is the prediction and, hence, the communication of the transport means charging availability with respect to both the traffic flows and the traffic conditions.

Whether the publications analyzed so far have in common the fact that the focus on private road users is conveyed mainly by people vehicles, [105] applied Markov models to smart city transport operations demonstrating how the sharing of the transport load in a smart city context can relevantly contribute to satisfying the mobility demand. According to the authors, the optimization of the individual transportation supply contributes to a more collaborative system leading to a more efficient, hence improved, transportation of loads.

4.2.7. *Vulnerable mobility*

Walkability in smart cities and the influence that such feature has on pedestrians' behavior in traffic is analyzed by [106]. They present and discuss the results of a macro-simulation study that aims at determining the walkability effect on people's behavior, by collecting GPS data at regular time intervals and categorizing them based on the interaction among different traffic modes. A Bayesian model is then applied to inspect the behavior of pedestrians, finally elaborating on related walkability indexes. Specifically targeting pedestrians, [107] proposed a solution for developing Big Data digital ecosystems like, e.g., the Smart Pedestrians Network, envisioning the use of Big Data generated from the combined use of GIS, IoT, crowdsourcing, and social networks. The final goal is boosting the walkability of the urban areas, fostering the environmental sustainability of the smart cities and, hence, their livability. In the specific ITS context operated by the Melbourne Smart City Office (Australia) [108] identify pedestrians mobility networks with the related hot spots and, through different data on parking lots, mobility, land use, environment, etc. they envisage the smart urban planning, e.g., the choice of future locations of traffic lights, streets configurations, huge infrastructural projects. Further developments target, for example, the embedding of social media data collected from smartphones and wearables.

[109] reviewed the use of Big Data with a specific focus on cyclists, specifically targeting the use of fitness apps and the information retrieved during several cycling initiatives, as well as by means of bike-sharing programs. This is done to fully exploit the interconnected potentialities from the combined use of Big Data in health and transportation fields. Also, [110] integrate the use of IoT and cloud computing to improve the safety of cyclists driving electric bicycles in urban areas context. In this regard, a specifically designed Raspberry system has been developed to enhance the continuous connection between the cyclists and the road infrastructure resulting in higher safety levels for cyclists. The authors mention how the complex data management system of the developed prototype can affect the performance of the system.

4.2.8. *Summary of the review findings*

Table 1 depicts a summary of the findings from the literature review about the publications presented and discussed so far in Section 4.

In Fig. 6 we resume the analytical work presented and discussed in Section 4 aiming to offer a roadmap for further research directions and analyses.

5. Discussion

This paper brings to light the significant effort dedicated by the existing literature on the subject toward the analytical and methodological development of (mainly Spatial) Big Data solutions in the context of smart cities mobility. This massive effort acts partially to the detriment of the effective deployment of such solutions to improve mobility in a sustainable way as well as of the evaluation of the proposed solutions' effectiveness and efficiency in practice and in real life

Table 1
Summary of the findings about the publications presented and discussed.

	Frequency
Publication type	
Journal paper	62
Conference paper	13
Chapter/book	8
Other	0
Open Access	
Yes	30
No	53
Year	
2010	0
2011	0
2012	0
2013	0
2014	4
2015	5
2016	20
2017	11
2018	16
2019	8
2020	3
2021	3
2022	13
Subject (related to the Smart Mobility Architecture model*)	
Data analytics	48
Private users	24
Public users	5
Vulnerable users	6

* For further details, see Fig. 5.

applications. In addition, there is a shortage of considerations about the impact of such solutions. The existing literature proposes theoretical and practical solutions to several problems related to mobility in the context of smart cities built on the use of Big Data. The profitability of Big Data-driven solutions is often evident and to some extent well-consolidated. Nevertheless, to what extent such solutions are effective in delivering an incremented level of sustainable existence to smart cities' inhabitants, this is less persistent.

Fig. 6 describes the depicted challenges from the present work which are supported, both from the articles analyzed in Section 4 and, also, from other researchers. For instance, the challenge of data openness is supported by [111] whose work presents how emerging technologies and open data from public authorities and other organizations can positively impact the decision process of any stakeholder related to mobility. Furthermore, a system dynamic approach to support decisionmakers, developed by [37], offers a take home message that lies in the *"validation of reliability"* (across different scenarios) of the model proposed in real-life cases. In addition, a potential further improvement is stressed to be the integration of dynamic databases in the solution(s) supporting data processing in real-time. The review from [112] stresses how transportation requires to gain designed, standardized approaches to collect, structure, store, and analyze Big Data through a conjoint effort from research and industry [113] use open-source environment data to map the most environment-friendly journey for a road user by envisioning the implementation of weather and traffic data processing, together with environmental data. Future research should be directed to the effective implementation of these combined data solutions.

[114] Define the opportunities and challenges regarding smart mobility in order to enhance more sustainable urban areas. Their findings envision the achievement of sustainable urban areas through the combined use of smart technologies and a cultural transition from city inhabitants. The latter, above all, envisions the sensibilization of citizens regarding the environmental impact of their mobility choice. Such sensibilization, in turn, it reflects in a shift from the use of private mobility towards shared mobility alternatives. The same authors recognize the importance of Big Data use to ameliorate the sustainability of cities. In the same context, [115] present how data-driven technologies impact the sustainability of smart cities. Data-driven solutions can be applied in several domains of the smart cities having potential

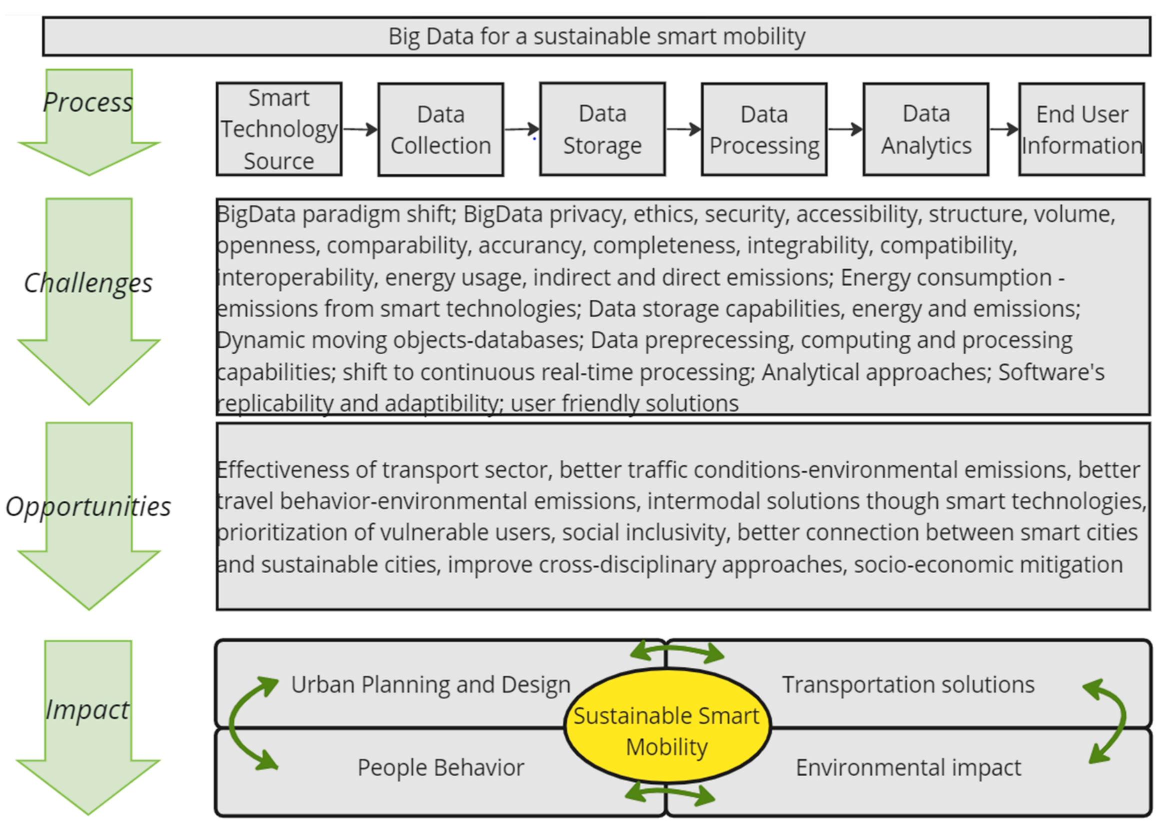

Fig. 6. Conceptual model of research findings for future research direction and analyses.

impacts on several systems (economic, environmental, and social). The same discuss that such data-driven smart cities are depended on both the technological advancement and the urban context. Interestingly, the authors point out a lack of "*data-driven smart cities indexes*" which can be a valuable tool to characterize the level of sustainability of a smart city based on its domains (e.g., in terms of mobility and its link with the environmental emissions).

Knowledge-based assessments from the mobility users are fundamental to support the mentioned cultural transition, sustaining more environment-friendly mobility decisions. In agreement with the mentioned literature, it is observed a shortage of well-established use of indicators to evaluate the real-life impact of the proposed solutions and/or the lack of cohesive frameworks within which it is possible to compare different choice scenarios operated by mobility users.

5.1. Theoretical implications

From the point of view of the methods, the advances promoted in the field are to a large extent well-established and consolidated. Among the others, solutions promoting the implementation of artificial intelligence systems and those involving diverse analytical tools in real-time are envisioned to play a major role in urban transportation planning and its management. However, the main challenge lies in the effort that must be done in ascribing all the different approaches in one cohesive framework that could be i) effective in solving the issues, ii) open to all the actors involved in the system (policymakers, planners, direct users, businesses, etc.), iii) focused on not disregarding the needs related to the information timeliness and accessibility.

To some extent, such considerations are in line with the findings of [34,37] and, despite the fact that the subject treated here is peculiarly depicted as fast-changing and extremely innovative, also to what envisaged by [116], almost 10 years ago. This hint, alone, it can support the need for effectively addressing the core challenge(s) offered by the topic, instead of providing innovative solutions that target peculiar issues in a very fragmented way.

In addition, the methodological advances proposed should be framed in a more general context, where not only the specific problem is considered, but, as well, there is space (and legitimacy) for the evaluation of the real impact(s) provided (if any) in terms of smart cities' sustainability. Indeed, this is in contrast with the findings of, e.g., [23,41,42] that discussed the use of Big Data from several perspectives, but disregarding all the matters linked to sustainability (here intended both in environmental and human terms).

One, residual, huge challenge that is yet linked with a strictly theoretical concern is offered, perhaps, by what [38] called the need for different Big Data(sets) to talk to each other being, hence, interconnected in real-time and openly accessible/accessed to manipulation and use at different scales and by different users.

At the edge of the future developments concerning the IoT systems used in transportation and the whole ITS, lately emerged the possibility of integrating Big Data and machine learning/data mining techniques with the blockchain technology. The latter has been widely addressed as the methodology that can play a crucial role in facilitating the transition from a linear to a circular economy and, as well, which can enable sustainability [117]. The integration of blockchain technology in ITS and in the framework of vehicle routing is increasingly considered the answer to the main pending research gap existing in the field [118]. The application of blockchain technology to solve ITS-related problems is largely envisaged with the purpose of solving some well-known Big Data issues, such as security problems and privacy disclosure risks [119]. In this regard, for example, [120] proposed to integrate machine learning and blockchain within the same hierarchical trajectory anomaly detection scheme for ITS. The proposal is to integrate blockchain technology to allow both on-chain and off-chain coordinated data access and help in building the trajectory anomaly detection model. This is done by limiting the risk of disclosure but, also, by reducing the uncertainty, infinity, and time-varying evolution issues (as well as the sparsity of data) related to Big Data use. Furthermore, [121] proposed a scheme for creating a data security aggregation protocol that, by allowing the establishment of an end-to-end encryption mechanism, protects the users from the risk of information leak and that of unwilling data mining when their vehicles are part of IoT/IoV systems.

5.2. Policy implications

The analytical solutions discussed so far can support activities such as monitoring, prediction, and planning, both in terms of public and private mobility. Nevertheless, embedding sustainability in smart mobility does not mean just developing innovative technologies, but, also, integrating the already existing solutions (grouped and systematically organized) with an environmental and human breath. In addition, there is the major need for quantitatively assessing the impact(s) offered by such solutions on smart cities' sustainability.

Form the political point of view, this can be done, for example, by starting to offer more responsible environmental options to direct users. Indeed, to the best of our knowledge, as well as from the results emerging from [48–55], mobility solution chosen by the city inhabitants are mainly based on the convenience of the solution *per se* with respect to the specific cases in terms of comfort (e.g., private vehicle rather than public transport), time, and economic costs (e.g., shorter travel time, cheaper travel solution) without assessing the environmental impact of the choice. The environmental issue is (partially) considered as a core concern only by [49,50], who approach the monitoring challenge in terms of eco-routing, or by considering the possibility to expressly monitor the "*performance of emissions*". Please, note that limitedly to [57,101], there is an explicit consideration of, e.g., the lowering of CO_2 emissions because of the innovation proposed, or in terms of the impact of electric (private) vehicles' usage. This is in contrast with the urgent call for sensibilization regarding the sustainable matters, and, as long as the mobility user is not recognizing sustainability as convenient for his mobility decision, the mobility will encounter difficulties into its sustainable development no matter how smart it is. Albeit this might be seen as a substantial challenge to achieve, we suggest that it is possible to start attracting the mobility user into sustainability, e.g., by using Big Data for increasing the attractiveness of the environmental aspect and the related convenience of the chosen solution, offering alternative scenarios of choice with the related (estimated) negative externalities produced and/or the positive ones providable.

Yet in the same line of the environmental sustainability, the problem of energy-saving through a Big Data-oriented approach to mobility is presented and discussed in [39,73], although this is done by stressing the economic sustainability. Hence, the point of view is yet linked to the convenience of the solution *per se.* The mobility of humans is then more largely discussed in terms of trajectories-learning and increase of efficiency for the end user (see, e.g., [66–72]). The matter of the overall livability of the smart city occupies a residual place in the discussion with, e.g., [65] debating about the quicker access to health, vital services like hospitals as being granted and/or ameliorated by means of more efficient people's trajectories.

Residual aspects are then related to the possibility for accounting for economic sustainability when the resources of policymakers and policy-planners of, e.g., public transports are scarce, as per the case described in [82]. In these cases, the 'livability' of smart cities as per the inhabitants' perceptions can increase only by means of a punctual and sharp use of the resources (both financial and material) but, not by disregarding the ability of policymakers and urban planners to gain credit in transparent and inclusive ways [84].

The latter point raises, above-all, in relation to the 'weakest' actors of the system to which a very marginal attention has been given so far (as proved by the solely work addressing the challenge of smart cities' walkability, [107]), even though they can strongly and decisively support the transition towards the sustainable path.

6. Conclusions

This paper presents and discusses the results from a hierarchical, Systematic Literature Review that covers the years 2010–2022, trying to answer the following research question: *"Which role has Big Data in fostering the smart city mobility?"*. By employing the SLR method, we retrieved from two web repositories (Google Scholar and Scopus) the papers published in relation to the use of Big Data for all the purposes linked to smart cities mobility. 2022 articles were initially retrieved. These, by trimming the results according to the SLR method, were reduced to 83 articles.

One limitation of the present study could lie in the exclusion of the gray literature on the subject. For the sake of brevity, as well as due to the shortage of potential robustness checks, our choice was oriented towards not considering the non-scientific records here. Another limitation can be due to the decision to not consider a third, widespread repository of scientific publications (e.g., Web of Science). In this regard, however, we are sufficiently confident that the records obtained here would have been overlapping among all the repositories (i.e., there is a tiny probability of publications not being covered by the two web repositories that have been scraped; for example, due to non-indexing).

Our results hint at the fact that bringing sustainability into the decisions of the mobility users means, for example, quantifying and evaluating the environmental impact(s) of the mobility decision, and, also, giving to direct users the proper tools for deciding by themselves how to reduce or mitigate the negative externalities deriving from their decisions. Furthermore, it is envisaged raising the knowledge and self-awareness of the direct users in terms of the potential positive externalities that, instead, can be produced by alternative choices.

Environmental impact indicators can be configured and implemented in the existing DaaS and SaaS solutions. Parameters such as CO_2 and NO_x emissions, and energy consumption, can be integrated into the already existing DaaS and SaaS giving as output the level of the environmental impact needed for the assessment of the sustainable impact of the mobility choice. This will give the mobility user the self-awareness needed to responsibly decide his mobility mode. To do so, the already existing DaaS and SaaS can be adopted and properly integrated with mobility platforms aiming to select the most sustainable intermodal mobility solution. Certainly, we do not disregard the relevance of challenges related to, e.g., the fact that merging mobility and environmental information in the same DaaS and SaaS means solving privacy, openness, and dynamics of databases issues. These, evidently, yet represent limitations to the sustainable transformation that should be operated in this context.

Some possible research directions for future works and further developments lie in the following:

- Researchers should focus on how to integrate the already (quite a few) existing solutions into more comprehensive mobility solutions, e.g., by including the environmental concerns into the platform.
- Possible, novel research questions to be addressed are *"How sustainability can be accounted as a convenience, or responsibility, for the citizen while deciding the smart mobility mode?"*. Also, *"How the environmental pollution affects the mobility user's travel decision in smart mobility?"*.
- Straightforwardly, questions like the aforementioned ones could be the key to sensibilize the mobility users towards the decision of their mobility mode. However, they can't be replied to without a co-joint effort between public, private, and research stakeholders.

The ongoing discussion on the sustainability of smart mobility must expose the citizens to the awareness of how they are actively involved in making the (smart) city more livable and sustainable. Smart mobility cannot be only framed into the use of innovative technologies and data services. Instead, such technologies must be deployed as a tool for bringing self-awareness to the mobility users about their environmental impact while choosing a (smart) mobility solution, possibly favoring the ability to walk, cycle, and scooter (as well as multimodality). At the same time, sustainability must be experienced as an additional "value" for the users, while choosing a given solution. This process of experiencing the convenience, and how sustainable the selected choice is, might bring mobility users closer to the concerns of sustainable development goals.

Author contributions

The authors confirm to have contributed to the preparation of the manuscript as follows:

Conceptualization and Design: H.G.; Methodology: R.D'A.; Formal Analysis: R.D'A. and H.G.; Investigation: R.D'A. and H.G.; Resources and Data Curation: R.D'A. and H.G.; Writing – abstract: R.D'A.; Section 1 – introduction: H.G.; Section 2 – big data: R.D'A.; Section 3 – review method and SLR results: R.D'A.; Section 4 – sustainable framework from (smart) city and (smart) mobility: H.G.; Section 5 – discussion: R.D'A. and H.G.; Section 6 – conclusions: R.D'A.; Review and Editing: R.D'A.; Visualization: R.D'A. and H.G. Final revision: R.D'A. and H.G.

Appendices/Supplemental material

None.

Declaration of Competing 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.

Data availability

Data will be made available on request.

References

[1] United 4 Smart Sustainable Cities, Digital Solutions for Integrated City Management & Use Cases – A U4SSC Deliverable on City Platforms, ITU, Geneva, 2021.

[2] Statista, Global Smart City Projects By Region 2022, Statista, 2022. www.statista.com/statistics/801167/projectofglobalsmartcitiesbyregion/. accessed August 12.

[3] European Commission, Smart Cities Marketplace 2022, European Commission, 2022. https://smart-cities-marketplace.ec.europa.eu. accessed August 17.

[4] K. Steinert, R. Marom, P. Richard, G. Veiga, L. Witters, Making Cities Smart and Sustainable. The global Innovation Index, INSEAD, Fontainbleau, 2011.

[5] J.R. Ferrer, Barcelona's Smart City vision: an opportunity for transformation, Field Actions Sci. Rep. 16 (2017) 70–75.

[6] European Commission, GREEN PAPER On The lmpact of Transport on the Environment, European Commission, 1992.

[7] P. Næss, T. Næss, A. Strand, The Challenge of Sustainable Mobility in Urban Planning and Development in Oslo Metropolitan Area, Institute of Transport Economics, Oslo, 2009.

[8] European Commission, Mainstreaming Sustainable Development Into EU policies: 2009 Review of the European Union Strategy for Sustainable Development, European Commission, 2009.

[9] K. Lu, B. Han, X. Zhou, Smart urban transit systems: from integrated framework to interdisciplinary perspective, Urban Rail Transit 4 (2018) 49–67, https://doi.org/10.1007/s40864-018-0080-x.

[10] L. Yang, K.H. van Dam, A. Majumdar, B. Anvari, W.Y. Ochieng, L. Zhang, Integrated design of transport infrastructure and public spaces considering human behavior: a review of state-of-the-art methods and tools, Front. Archit. Res. 8 (2019) 429–453, https://doi.org/10.1016/j.foar.2019.08.003.

[11] V. Albino, U. Berardi, R.M. Dangelico, Smart cities: definitions, dimensions, performance, and initiatives, J. Urban Technol. 22 (2015) 3–21, https://doi.org/10.1080/10630732.2014.942092.

[12] C. Harrison, B. Eckman, R. Hamilton, P. Hartswick, J. Kalagnanam, J. Paraszczak, et al., Foundations for Smarter Cities, IBM J. Res. Dev. 54 (2010) 1–16, https://doi.org/10.1147/JRD.2010.2048257.

[13] C. Lim, K.J. Kim, P.P. Maglio, Smart cities with big data: reference models, challenges, and considerations, Cities 82 (2018) 86–99, https://doi.org/10.1016/j.cities.2018.04.011.

[14] B. Jiang, Z. Ren, Geographic space as a living structure for predicting human activities using big data, Int. J. Geogr. Inf. Sci. 33 (2019) 764–779, https://doi.org/10.1080/13658816.2018.1427754.
[15] M. Batty, Big data, smart cities and city planning, Dialogues Hum. Geogr. 3 (2013) 274–279, https://doi.org/10.1177/2043820613513390.
[16] I. Vilajosana, J. Llosa, B. Martinez, M. Domingo-Prieto, A. Angles, X. Vilajosana, Bootstrapping smart cities through a self-sustainable model based on big data flows, IEEE Commun. Mag. 51 (2013) 128–134, https://doi.org/10.1109/MCOM.2013.6525605.
[17] R. Kitchin, The real-time city? Big data and smart urbanism, GeoJ. 79 (2014) 1–14, https://doi.org/10.1007/s10708-013-9516-8.
[18] E. Al Nuaimi, H. Al Neyadi, N. Mohamed, J Al-Jaroodi, Applications of big data to smart cities, J. Internet Serv. Appl. 6 (2015) 25, https://doi.org/10.1186/s13174-015-0041-5.
[19] V.M. Gabel, R. Home, M. Stolze, S. Birrer, B. Steinemann, U. Köpke, The influence of on-farm advice on beliefs and motivations for Swiss lowland farmers to implement ecological compensation areas on their farms, J. Agric. Educ. Ext. 24 (2018) 233–248, https://doi.org/10.1080/1389224X.2018.1428205.
[20] H. Gaber, A.M. Othman, A.H. Fahad, Future of connected autonomous vehicles in smart cities. Solving Urban Infrastructure Problems Using Smart City Technologies, Elsevier, 2021, pp. 599–611, https://doi.org/10.1016/B978-0-12-816816-5.00027-9.
[21] S. Porru, F.E. Misso, F.E. Pani, C. Repetto, Smart mobility and public transport: opportunities and challenges in rural and urban areas, J. Traffic Transp. Eng. 7 (2020) 88–97, https://doi.org/10.1016/j.jtte.2019.10.002 (English Edition).
[22] A. Khamis, Smart Mobility: Technology Enablers. Smart Mobility, Apress, Berkeley, 2021.
[23] L. Zhu, F.R. Yu, Y. Wang, B. Ning, T. Tang, Big data analytics in intelligent transportation systems: a survey, IEEE Trans Intell Transport Syst 20 (2019) 383–398, https://doi.org/10.1109/TITS.2018.2815678.
[24] A. Daniel, A. Paul, A. Ahmad, Near real-time big data analysis on vehicular networks, in: Proceedings of the 2015 International Conference on Soft-Computing and Networks Security (ICSNS), Coimbatore, India, 2015, pp. 1–7, https://doi.org/10.1109/ICSNS.2015.7292404. IEEE.
[25] A.I. Torre-Bastida, J. Del Ser, I. Laña, M. Ilardia, M.N. Bilbao, S Campos-Cordobés, Big data for transportation and mobility: recent advances, trends and challenges, IET Intel. Transport Syst. 12 (2018) 742–755, https://doi.org/10.1049/iet-its.2018.5188.
[26] C.J.Q. Farmer, A. Pozdnoukhov, Building streaming GIScience from context, theory, and intelligence, in: Proceedings of the Workshop on GIScience in the Big Data Age, Columbus, 2012, pp. 5–10.
[27] R. Kitchin, G. McArdle, What makes big data, big data? exploring the ontological characteristics of 26 datasets, Big Data Soc 3 (2016), 205395171663113, https://doi.org/10.1177/2053951716631130.
[28] R. Herschel, V.M. Miori, Ethics & big data, Technol. Soc. 49 (2017) 31–36, https://doi.org/10.1016/j.techsoc.2017.03.003.
[29] M. Chen, S. Mao, Y. Liu, Big data: a survey, Mobile Netw Appl 19 (2014) 171–209, https://doi.org/10.1007/s11036-013-0489-0.
[30] Statista, Forecast Revenue Big Data Market Worldwide 2011-2027 2022, Statista, 2022. www.statista.com/statistics/254266/globalbigdata-market-forecast/. accessed July 27.
[31] M. MARJANI, F. NASARUDDIN, A. GANI, A. KARIM, I.A.T. HASHEM, A. SIDDIQA, et al., Big IoT data analytics: architecture, opportunities, and open research challenges, IEEE Access 5 (2017) 5247–5261, https://doi.org/10.1109/ACCESS.2017.2689040.
[32] Srinivasa S., Bhatnagar V., editors. Big Data Analytics: First International Conference, BDA 2012, New Delhi, India, December 24-26, 2012. Proceedings. vol. 7678. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. 10.1007/978-3-642-35542-4.
[33] J. Qiu, Q. Wu, G. Ding, Y. Xu, S. Feng, A survey of machine learning for big data processing, EURASIP J. Adv. Signal Process. 2016 (2016) 67, https://doi.org/10.1186/s13634-016-0355-x.
[34] S.E. Bibri, Big Data Science and Analytics for Smart Sustainable Urbanism: Unprecedented Paradigmatic Shifts and Practical Advancements, Springer International Publishing, Cham, 2019, https://doi.org/10.1007/978-3-030-17312-8.
[35] I. Anagnostopoulos, S. Zeadally, E. Exposito, Handling big data: research challenges and future directions, J Supercomput 72 (2016) 1494–1516, https://doi.org/10.1007/s11227-016-1677-z.
[36] D. Gough, S. Oliver, J. Thomas, An Introduction to Systematic Reviews, SAGE, 2017.
[37] P. Del Vecchio, G. Secundo, Y. Maruccia, G. Passiante, A system dynamic approach for the smart mobility of people: implications in the age of big data, Technol. Forecast. Soc. Change 149 (2019), 119771, https://doi.org/10.1016/j.techfore.2019.119771.
[38] C. Anda, A. Erath, P.J. Fourie, Transport modelling in the age of big data, Int. J. Urban Sci. 21 (2017) 19–42, https://doi.org/10.1080/12265934.2017.1281150.
[39] S. Giest, Big data analytics for mitigating carbon emissions in smart cities: opportunities and challenges, Eur. Plan. Studies 25 (2017) 941–957, https://doi.org/10.1080/09654313.2017.1294149.
[40] D. Milne, D. Watling, Big data and understanding change in the context of planning transport systems, J. Transp. Geogr. 76 (2019) 235–244, https://doi.org/10.1016/j.jtrangeo.2017.11.004.
[41] N. Pelekis, Y. Theodoridis, Mobility Data Management and Exploration, Springer New York, New York, NY, 2014, https://doi.org/10.1007/978-1-4939-0392-4.
[42] S. Shukla, K. Balachandran, V.S. Sumitha, A framework for smart transportation using Big Data, in: Proceedings of the 2016 International Conference on ICT in Business Industry & Government (ICTBIG), Indore, India, IEEE, 2016, pp. 1–3, https://doi.org/10.1109/ICTBIG.2016.7892720.
[43] Vlahogianni E.I. Computational intelligence and optimization for transportation big data: challenges and opportunities. In: Lagaros N.D., Papadrakakis M., editors. Engineering and Applied Sciences Optimization, vol. 38, Cham: Springer International Publishing; 2015, p. 107–28. 10.1007/978-3-319-18320-6_7.
[44] K. Zhao, S. Tarkoma, S. Liu, H. Vo, Urban human mobility data mining: an overview, in: Proceedings of the 2016 IEEE International Conference on Big Data (Big Data), Washington DC,USA, 2016, pp. 1911–1920, https://doi.org/10.1109/BigData.2016.7840811. IEEE;.
[45] X. Zheng, W. Chen, P. Wang, D. Shen, S. Chen, X. Wang, et al., Big data for social transportation, IEEE Trans. Intell. Transport Syst 17 (2016) 620–630, https://doi.org/10.1109/TITS.2015.2480157.
[46] X. Wang, Z. Li, Integrated platform for smart traffic big data, in: Proceedings of the 2016 International Conference on Logistics, Informatics and Service Sciences (LISS), Sydney, Australia, IEEE, 2016, pp. 1–6, https://doi.org/10.1109/LISS.2016.7854592.
[47] S. Fiore, D. Elia, C.E. Pires, D.G. Mestre, C. Cappiello, M. Vitali, et al., An Integrated big and fast data analytics platform for smart urban transportation management, IEEE Access 7 (2019) 117652–117677, https://doi.org/10.1109/ACCESS.2019.2936941.
[48] P. Rizwan, K. Suresh, M.R. Babu, Real-time smart traffic management system for smart cities by using Internet of Things and big data, in: Proceedings of the 2016 International Conference on Emerging Technological Trends (ICETT), Kollam, India, IEEE, 2016, pp. 1–7, https://doi.org/10.1109/ICETT.2016.7873660.
[49] A. Saroj, S. Roy, A. Guin, M. Hunter, R. Fujimoto, Smart city real-time data-driven transportation simulation, in: Proceedings of the 2018 Winter Simulation Conference (WSC), Gothenburg, Sweden, 2018, pp. 857–868, https://doi.org/10.1109/WSC.2018.8632198. IEEE.
[50] C. Wang, X. Li, X. Zhou, A. Wang, N. Nedjah, Soft computing in big data intelligent transportation systems, Appl. Soft Comput. 38 (2016) 1099–1108, https://doi.org/10.1016/j.asoc.2015.06.006.
[51] S. Amini, I. Gerostathopoulos, C. Prehofer, Big data analytics architecture for real-time traffic control, in: Proceedings of the 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), Naples, Italy, 2017, pp. 710–715, https://doi.org/10.1109/MTITS.2017.8005605. IEEE.
[52] B. Jan, H. Farman, M. Khan, M. Talha, I.U. Din, Designing a smart transportation system: an internet of things and big data approach, IEEE Wireless Commun. 26 (2019) 73–79, https://doi.org/10.1109/MWC.2019.1800512.
[53] M.M. Rathore, A. Ahmad, A. Paul, U.K. Thikshaja, Exploiting real-time big data to empower smart transportation using big graphs, in: Proceedings of the 2016 IEEE Region 10 Symposium (TENSYMP), Bali, Indonesia, 2016, pp. 135–139, https://doi.org/10.1109/TENCONSpring.2016.7519392. IEEE.
[54] Z. Liu, Z. Li, K. Wu, M. Li, Urban traffic prediction from mobility data using deep learning, IEEE Netw. 32 (2018) 40–46, https://doi.org/10.1109/MNET.2018.1700411.
[55] G. Manogaran, T.N. Nguyen, Displacement-aware service endowment scheme for improving intelligent transportation systems data exchange, IEEE Trans. Intell. Transp. Syst. 23 (2022) 22467–22477, https://doi.org/10.1109/TITS.2021.3078753.
[56] Y. Liu, Big data technology and its analysis of application in urban intelligent transportation system, in: Proceedings of the 2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), Xiamen, China, 2018, pp. 17–19, https://doi.org/10.1109/ICITBS.2018.00012. IEEE.
[57] N. Cárdenas-Benítez, R. Aquino-Santos, P. Magaña-Espinoza, J. Aguilar-Velazco, A. Edwards-Block, A. Medina Cass, Traffic congestion detection system through connected vehicles and big data, Sensors 16 (2016) 599, https://doi.org/10.3390/s16050599.
[58] W.Y.H. Adoni, N.B. Aoun, T. Nahhal, M. Krichen, M.Y. Alzahrani, F.K. Mutombo, A scalable big data framework for real-time traffic monitoring system, J. Comput. Sci. 18 (2022) 801–810, https://doi.org/10.3844/jcssp.2022.801.810.
[59] T. Devi, K. Alice, N Deepa, Traffic management in smart cities using support vector machine for predicting the accuracy during peak traffic conditions, Mater. Today Proc. 62 (2022) 4980–4984, https://doi.org/10.1016/j.matpr.2022.03.722.
[60] F.P. Diez, J.C. Sinca, D.R. Vallès, J.M. Campos Cacheda, Evaluation of transport events with the use of big data, artificial intelligence and augmented reality techniques, Transp. Res. Procedia 58 (2021) 173–180, https://doi.org/10.1016/j.trpro.2021.11.024.
[61] Y. Karouani, Z. Elhoussaine, M. Ben Ahmed, A.A. Boudhir, Toward an intelligent traffic management based on big data for smart city, editors, in: Innovations in Smart Cities and Applications, 37, Springer International Publishing, Cham, 2018, pp. 502–514, https://doi.org/10.1007/978-3-319-74500-8_47.
[62] M. Humayun, S. Afsar, M.F. Almufareh, N.Z. Jhanjhi, M. AlSuwailem, Smart Traffic management system for metropolitan cities of kingdom using cutting edge technologies, J. Adv. Transp. 2022 (2022) 1–13, https://doi.org/10.1155/2022/4687319.
[63] F. Xu, Y. Lin, J. Huang, D. Wu, H. Shi, J. Song, et al., Big data driven mobile traffic understanding and forecasting: a time series approach, IEEE Trans. Serv. Comput. 9 (2016) 796–805, https://doi.org/10.1109/TSC.2016.2599878.
[64] Y. Liu, X. Weng, J. Wan, X. Yue, H. Song, A.V. Vasilakos, Exploring data validity in transportation systems for smart Cities, IEEE Commun. Mag. 55 (2017) 26–33, https://doi.org/10.1109/MCOM.2017.1600240.

[65] C.J. Pettit, S. Zarpelon Leao, O. Lock, M. Ng, J Reades, Big data: the engine to future cities—a reflective case study in urban transport, Sustainability 14 (2022) 1727, https://doi.org/10.3390/su14031727.

[66] C. Chen, J. Ma, Y. Susilo, Y. Liu, M. Wang, The promises of big data and small data for travel behavior (aka human mobility) analysis, Transp. Res. C Emerg. Technol. 68 (2016) 285–299, https://doi.org/10.1016/j.trc.2016.04.005.

[67] Q. Wang, J.E. Taylor, Process map for urban-human mobility and civil infrastructure data collection using geosocial networking platforms, J. Comput. Civ. Eng. 30 (2016), 04015004, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000469.

[68] G. Xu, S. Gao, M. Daneshmand, C. Wang, Y. Liu, A survey for mobility big data analytics for geolocation prediction, IEEE Wireless Commun. 24 (2017) 111–119, https://doi.org/10.1109/MWC.2016.1500131WC.

[69] M. Itoh, D. Yokoyama, M. Toyoda, Y. Tomita, S. Kawamura, M. Kitsuregawa, Visual fusion of mega-city big data: an application to traffic and tweets data analysis of metro passengers, in: Proceedings of the 2014 IEEE International Conference on Big Data (Big Data), Washington, DC, USA, IEEE, 2014, pp. 431–440, https://doi.org/10.1109/BigData.2014.7004260.

[70] N. Zhang, H. Chen, X. Chen, J. Chen, J. Cao, K. Mao, J. Wu, A. Lendasse, ELM meets urban computing: ensemble urban data for smart city application, editors, in: Proceedings of the ELM-2015 Volume 1, 6, Springer International Publishing, 2016, pp. 51–63, https://doi.org/10.1007/978-3-319-28397-5_5.

[71] Y. Qiao, Y. Cheng, J. Yang, J. Liu, N. Kato, A mobility analytical framework for big mobile data in densely populated area, IEEE Trans. Veh. Technol. 66 (2017) 1443–1455, https://doi.org/10.1109/TVT.2016.2553182.

[72] G. Di Lorenzo, M. Sbodio, F. Calabrese, M. Berlingerio, F. Pinelli, R. Nair, AllAboard: visual exploration of cellphone mobility data to optimise public transport, IEEE Trans. Vis. Comput. Graph. 22 (2016) 1036–1050, https://doi.org/10.1109/TVCG.2015.2440259.

[73] F. Pinelli, R. Nair, F. Calabrese, M. Berlingerio, G. Di Lorenzo, M.L. Sbodio, Data-driven transit network design from mobile phone trajectories, IEEE Trans.Intell. Transp. Syst. 17 (2016) 1724–1733, https://doi.org/10.1109/TITS.2015.2496783.

[74] D. Xu, H. Wang, Y. Li, S. Tarkoma, D. Jin, P. Hui, IoT vs. human: a comparison of mobility, IEEE Trans.Mobile Comput. 21 (2022) 1257–1273, https://doi.org/10.1109/TMC.2020.3019988.

[75] J. Wu, S. Guo, J. Li, D. Zeng, Big data meet green challenges: big data toward green applications, IEEE Syst. J. 10 (2016) 888–900, https://doi.org/10.1109/JSYST.2016.2550530.

[76] T. Zaree, A.R. Honarvar, Improvement of air pollution prediction in a smart city and its correlation with weather conditions using metrological big data, Turk. J. Elec. Eng. Comput. Sci. 26 (2018), https://doi.org/10.3906/elk-1707-99.

[77] S. Zhang, T. Niu, Y. Wu, K.M. Zhang, T.J. Wallington, Q. Xie, et al., Fine-grained vehicle emission management using intelligent transportation system data, Environ. Pollut. 241 (2018) 1027–1037, https://doi.org/10.1016/j.envpol.2018.06.016.

[78] N. van Oort, D. Sparing, T. Brands, R.M.P. Goverde, Data driven improvements in public transport: the Dutch example, Public Transp. 7 (2015) 369–389, https://doi.org/10.1007/s12469-015-0114-7.

[79] Z. Huang, X. Ling, P. Wang, F. Zhang, Y. Mao, T. Lin, et al., Modeling real-time human mobility based on mobile phone and transportation data fusion, Transp. Res. C Emerg. Technol. 96 (2018) 251–269, https://doi.org/10.1016/j.trc.2018.09.016.

[80] H. Poonawala, V. Kolar, S. Blandin, L. Wynter, S. Sahu, Singapore in motion: insights on public transport service level through farecard and mobile data analytics, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco California USA, ACM, 2016, pp. 589–598, https://doi.org/10.1145/2939672.2939723.

[81] S. Tao, J. Corcoran, I. Mateo-Babiano, D. Rohde, Exploring bus rapid transit passenger travel behaviour using big data, Appl. Geogr. 53 (2014) 90–104, https://doi.org/10.1016/j.apgeog.2014.06.008.

[82] Y. Wang, S. Ram, F. Currim, E. Dantas, L.A. Saboia, A big data approach for smart transportation management on bus network, in: Proceedings of the 2016 IEEE International Smart Cities Conference (ISC2), Trento, Italy, IEEE, 2016, pp. 1–6, https://doi.org/10.1109/ISC2.2016.7580839.

[83] F.C. Pereira, F. Rodrigues, M. Ben-Akiva, Using data from the web to predict public transport arrivals under special events scenarios, J. Intell. Transp. Syst. 19 (2015) 273–288, https://doi.org/10.1080/15472450.2013.868284.

[84] C. Lemonde, E. Arsenio, R. Henriques, Integrative analysis of multimodal traffic data: addressing open challenges using big data analytics in the city of Lisbon, Eur. Transp. Res. Rev. 13 (2021) 64, https://doi.org/10.1186/s12544-021-00520-3.

[85] S. Cerqueira, E. Arsenio, R. Henriques, Inference of dynamic origin–destination matrices with trip and transfer status from individual smart card data, Eur. Transp. Res. Rev. 14 (2022) 42, https://doi.org/10.1186/s12544-022-00562-1.

[86] S. Schatzinger, C.Y.R. Lim, Taxi of the future: big data analysis as a framework for future urban fleets in smart cities, Smart Sustain. Plan. Cities Reg. (2016) 83–98.

[87] X. Zhang, Y. Xu, W. Tu, C. Ratti, Do different datasets tell the same story about urban mobility — a comparative study of public transit and taxi usage, J. Transp. Geogr. 70 (2018) 78–90, https://doi.org/10.1016/j.jtrangeo.2018.05.002.

[88] Z. Zhou, W. Dou, G. Jia, C. Hu, X. Xu, X. Wu, et al., A method for real-time trajectory monitoring to improve taxi service using GPS big data, Inform. Manag. 53 (2016) 964–977, https://doi.org/10.1016/j.im.2016.04.004.

[89] D. Zhang, T. He, S. Lin, S. Munir, J.A. Stankovic, Taxi-passenger-demand modeling based on big data from a roving sensor network, IEEE Trans. Big Data 3 (2017) 362–374, https://doi.org/10.1109/TBDATA.2016.2627224.

[90] W. Yuan, P. Deng, T. Taleb, J. Wan, C. Bi, An unlicensed taxi identification model based on big data analysis, IEEE Trans. Intell. Transp. Syst. 17 (2016) 1703–1713, https://doi.org/10.1109/TITS.2015.2498180.

[91] Cheng N., Lyu F., Chen J., Xu W., Zhou H., Zhang S., et al. Big Data Driven Vehicular Networks 2018.

[92] P. Bedi, V. Jindal, Use of big data technology in vehicular Ad-hoc networks, in: Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Delhi, India, IEEE, 2014, pp. 1677–1683, https://doi.org/10.1109/ICACCI.2014.6968352.

[93] W. Xu, H. Zhou, N. Cheng, F. Lyu, W. Shi, J. Chen, et al., Internet of vehicles in big data era, IEEE CAA J. Autom. Sin. 5 (2018) 19–35, https://doi.org/10.1109/JAS.2017.7510736.

[94] Q. Cui, Y. Wang, K.C. Chen, W. Ni, I.C. Lin, X. Tao, et al., Big data analytics and network calculus enabling intelligent management of autonomous vehicles in a smart city, IEEE Internet Things J. 6 (2019) 2021–2034, https://doi.org/10.1109/JIOT.2018.2872442.

[95] Malik AW Murk, I. Mahmood, N. Ahmed, Z Anwar, Big data in motion: a vehicle-assisted urban computing framework for smart cities, IEEE Access 7 (2019) 55951–55965, https://doi.org/10.1109/ACCESS.2019.2913150.

[96] Anthony B., Petersen S.A. A practice based exploration on electric mobility as a service in smart cities. In: Themistocleous M., Papadaki M., editors. Information Systems, Cham: Springer International Publishing; 2020, vol. 381, p. 3–17. 10.1007/978-3-030-44322-1_1.

[97] A. Bennett, Autonomous vehicle driving algorithms and smart m obility technologies in big data-driven transportation planning and engineering, Contemp. Read. Law Soc. Justice 13 (2021) 20–29.

[98] A.B.T. Sherif, K. Rabieh, M.M.E.A. Mahmoud, X Liang, Privacy-preserving ride sharing scheme for autonomous vehicles in big data era, IEEE Internet Things J. 4 (2017) 611–618, https://doi.org/10.1109/JIOT.2016.2569090.

[99] Alajali W., Zhou W. On-street car parking prediction in smart city: a multi-source data analysis in sensor-cloud Environment. Security, Privacy, and Anonymity in Computation, Communication, and Storage, 2017, p. 641–52.

[100] L. Hu, A. Liu, M. Xie, T. Wang, UAVs joint vehicles as data mules for fast codes dissemination for edge networking in Smart City, Peer to Peer Netw. Appl. 12 (2019) 1550–1574, https://doi.org/10.1007/s12083-019-00752-0.

[101] B. Li, M.C. Kisacikoglu, C. Liu, N. Singh, M. Erol-Kantarci, Big data analytics for electric vehicle integration in green smart cities, IEEE Commun. Mag. 55 (2017) 19–25, https://doi.org/10.1109/MCOM.2017.1700133.

[102] B. Anthony Jnr, S. Abbas Petersen, D. Ahlers, J Krogstie, Big data driven multi-tier architecture for electric mobility as a service in smart cities: a design science approach, IJESM 14 (2020) 1023–1047, https://doi.org/10.1108/IJESM-08-2019-0001.

[103] M. Hamdani, N. Sahli, N. Jabeur, N. Khezami, Agent-based approach for connected vehicles and smart road signs collaboration, Cai 41 (2022) 376–396, https://doi.org/10.31577/cai_2022_1_376.

[104] Y. Cao, H. Song, O. Kaiwartya, B. Zhou, Y. Zhuang, Y. Cao, et al., Mobile edge computing for big-data-enabled electric vehicle charging, IEEE Commun. Mag. 56 (2018) 150–156, https://doi.org/10.1109/MCOM.2018.1700210.

[105] R. Mehmood, R. Meriton, G. Graham, P. Hennelly, M. Kumar, Exploring the influence of big data on city transport operations: a Markovian approach, IJOPM 37 (2017) 75–104, https://doi.org/10.1108/IJOPM-03-2015-0179.

[106] Y. Yamagata, D. Murakami, Y. Wu, P.P.J. Yang, T. Yoshida, R. Binder, Big-data analysis for carbon emission reduction from cars: towards walkable green smart community, Energy Procedia 158 (2019) 4292–4297, https://doi.org/10.1016/j.egypro.2019.01.795.

[107] G. Papageorgiou, C. Petrakis, A. Maimaris, Developing a smart pedestrian network big data platform for municipal organizations, in: Proceedings of the 2018 2nd European Conference on Electrical Engineering and Computer Science (EECS), Bern, Switzerland, 2018, pp. 274–278, https://doi.org/10.1109/EECS.2018.00058. IEEE.

[108] E. Carter, P. Adam, D. Tsakis, S. Shaw, R. Watson, P. Ryan, Enhancing pedestrian mobility in Smart Cities using Big Data, J. Manag. Anal. 7 (2020) 173–188, https://doi.org/10.1080/23270012.2020.1741039.

[109] G. Romanillos, M. Zaltz Austwick, D. Ettema, J De Kruijf, Big data and cycling, Transp. Rev. 36 (2016) 114–133, https://doi.org/10.1080/01441647.2015.1084067.

[110] F. Martella, M. Fazio, G. Ciulla, R. Di Bernardo, A. Celesti, V. Lukaj, et al., An edge system for the safety of cyclists in the urban area, in: Proceedings of the 2022 IEEE International Smart Cities Conference (ISC2), Pafos, Cyprus, 2022, pp. 1–7, https://doi.org/10.1109/ISC255366.2022.9922454. IEEE.

[111] F.R. Soriano, J.J. Samper-Zapater, J.J. Martinez-Dura, R.V. Cirilo-Gimeno, J. Martinez Plume, Smart mobility trends: open data and other tools, IEEE Intell. Transp. Syst. Mag. 10 (2018) 6–16, https://doi.org/10.1109/MITS.2017.2743203.

[112] K. Soomro, M.N.M. Bhutta, Z. Khan, M.A. Tahir, Smart city big data analytics: an advanced review, WIREs Data Mining Knowl. Discov. 9 (2019), https://doi.org/10.1002/widm.1319.

[113] J. Zenkert, M. Dornhofer, C. Weber, C. Ngoukam, M. Fathi, Big data analytics in smart mobility: modeling and analysis of the Aarhus smart city dataset, in: Proceedings of the 2018 IEEE Industrial Cyber-Physical Systems (ICPS), St. Petersburg, 2018, pp. 363–368, https://doi.org/10.1109/ICPHYS.2018.8387685. IEEE.

[114] S. Paiva, M. Ahad, G. Tripathi, N. Feroz, G. Casalino, Enabling technologies for urban smart mobility: recent trends, opportunities and challenges, Sensors 21 (2021) 2143, https://doi.org/10.3390/s21062143.

[115] S.E. Bibri, J. Krogstie, The emerging data–driven smart city and its innovative applied solutions for sustainability: the cases of London and Barcelona, Energy Inform. 3 (2020) 5, https://doi.org/10.1186/s42162-020-00108-6.
[116] M.H. Tsou, Big data: techniques and technologies in geoinformatics, Ann. GIS 20 (2014) 295–296, https://doi.org/10.1080/19475683.2014.944934.
[117] A. Rejeb, S. Zailani, K. Rejeb, H. Treiblmaier, J.G. Keogh, Modeling enablers for blockchain adoption in the circular economy, Sustain. Futures 4 (2022), 100095, https://doi.org/10.1016/j.sftr.2022.100095.
[118] M. Salehi Sarbijan, J. Behnamian, Emerging research fields in vehicle routing problem: a short review, Arch. Computat. Methods Eng. 30 (2023) 2473–2491, https://doi.org/10.1007/s11831-022-09874-w.
[119] X. Guo, X. Guo, A research on blockchain technology: urban intelligent transportation systems in developing countries, IEEE Access 11 (2023) 40724–40740, https://doi.org/10.1109/ACCESS.2023.3270100.
[120] X. Wang, W. Liu, H. Lin, J. Hu, K. Kaur, M.S. Hossain, AI-empowered trajectory anomaly detection for intelligent transportation systems: a hierarchical federated learning approach, IEEE Trans. Intell. Transp. Syst. 24 (2023) 4631–4640, https://doi.org/10.1109/TITS.2022.3209903.
[121] C. Xu, H. Wu, H. Liu, W. Gu, Y. Li, D. Cao, Blockchain-oriented privacy protection of sensitive data in the internet of vehicles, IEEE Trans. Intell. Veh. 8 (2023) 1057–1067, https://doi.org/10.1109/TIV.2022.3164657.