Research Trends in the Application of Big Data in Smart Cities - A Literature Review

Abdelrahman Youssef[a], Petr Hajek[a], Lubica Hikkerova[b,*]

[a] Science and Research Centre, Faculty of Economics and Administration, University of Pardubice, Studentska 84, Pardubice, Czech Republic

[b] IPAG Business School, Paris, France

* corresponding author

e-mail: abdelrahmantaha.youssef@student.upce.cz, petr.hajek@upce.cz

This work was supported by a grant provided by the scientific research project of the Czech Sciences Foundation Grant No. 20-03037S and by the Student Grant Competition, grant No. SGS_2022_008.

Abstract

In this study, we review articles published in the last five years at the intersection of big data and the smart city context, focusing our analysis on a dataset of 192 articles. We use content analysis to identify recent trends and future research directions. The most influential research areas include using big data analytics, electric vehicle integration, and citizen-centric big data, while the most recent research area that emerged is green space. Our results highlight that big data can influence eight different dimensions of smart cities, with smart mobility attracting the most attention from researchers in the application of big data in smart cities. In contrast, smart people and smart economy are the least explored dimensions from this perspective.

Keywords: smart city, big data, application, literature review, content analysis.

1. Introduction

With the increasing number of devices such as mobile phones and various sensors connected to many products and infrastructure in daily use, and the interaction of many actors including humans, there is a massively increasing amount of data that can be called big data. Although smart cities represent an evolution and modern version of living spaces, big data plays a key role in how cities can be smarter and use their data to gain valuable insights that enable effective decision making. Big data is both the result and the building block of technological advancements that are the means to accelerate the development of smart cities. The use of big data can benefit smart cities in various sectors such as infrastructure, healthcare, transport, energy, and environment, among others (Batty, 2019; Kourtit et al., 2017).

Big data analytics for urban governance and planning purposes is crucial for the formation and operation of smart cities. It is assumed that as the accessibility of real-time high-frequency data increases, so does the ability of intelligent systems to detect patterns in the behaviour of elements of urban systems. The premise is that urban big data analysis can be instrumental in optimising city resources and solving long-term problems in cities (Kandt & Batty, 2021).

Despite the growing number of successful applications of big data in many cities around the world, understanding the use of big data for smart cities remains scarce in the current literature (Lim et al., 2018). Moreover, only few studies have provided an overview of the use of big data for smart cities, regardless of the application area. Notably, Hashem et al. (2016) defined the ecosystem of big data and the enabling technologies, highlighted how to standardize data management and data processing, and showed several scenarios of how smart cities can benefit from big data. Al Nuaimi et al. (2015) elaborated on the benefits and opportunities of big data for smart cities, specifying the enabling requirements to capture the benefits of big data and the associated challenges. Most recently, Karimi et al. (2021) systematically reviewed algorithm- and architecture-based approaches to the use of big data in smart cities.

However, prior studies on the use of big data in smart cities have tended to identify the benefits and challenges associated with the application of big data in smart cities while neglecting to identify key research trends and emerging research areas. Indeed, rapid changes in data architectures and artificial intelligence methods are having a serious effect on the use of big data in smart cities. Therefore, the aim of this work is to identify current research trends in this area. Specifically, this study adopts a content analysis approach to review recent research published between 2017 and 2022 related to the intersection of big data and smart cities. When analysing recent research efforts, research gaps can be highlighted and opportunities for future research are suggested. To further investigate the impact of big data on individual smart city dimensions, such as smart mobility and smart environment, we select a shortlist of relevant studies and discuss the potential of using big data to enhance each of those dimensions.

The importance of this study lies in providing an understanding of the current evolving trends of big data as an enabler for the development of smart cities, thus maximizing the full potential of smart city capabilities in the context of urban spaces. Several stakeholders such as urban planners, policy makers and technology leaders can benefit from this study to leverage big data to drive innovation and promote data-driven decisions. The applications and use cases of big data are increasing every day, as are the strategies for developing smart cities, so it is clear that city planners need to continuously adapt and evolve their strategies accordingly to keep up with the expectations of their citizens and to address emerging challenges and demands in the most effective way.

The remainder of this study is organized as follows. Section 2 presents related previous studies on the use of big data in smart cities. Section 3 describes the methodology used. Section 4 shows the results of the content analysis and typology of big data applications in terms of smart city dimensions. Section 5 discusses the results and suggests future research avenues, and Section 6 concludes.

2. Related Literature

Smart cities have specific characteristics that have information technology at their core. Smart cities rely on the development of human capital, which is a strong player in the development of the city. Efficient transport is needed to ensure smooth and fast mobility. Efficient governance and resource management are essential for sustainable development and economic growth. Infrastructure development and environmental protection are also essential to improve quality of life. The definitions of smart cities were summarized by Hajek et al. (2022), concluding that a smart city relies on ICT capability that improves resource efficiency, facilitates transportation, affects people, increases lining standards, promotes prosperity, and sustainability. The rise of smart cities is an outstanding illustration of the diffusion and utilization of big data analytics (Karimi et al., 2021).

Big data is often referred to as a large amount of data that is difficult to store and process with simple tools and requires advanced techniques; as descried by Jorgensen et al. (2014), big data processing requires high computing power, leveraging advanced techniques of machine learning and artificial intelligence to process these huge and complex datasets. Big data has three main characteristics defined by Laney (2001), often referred to as the 3Vs:

1. Volume: the data is in huge quantities, typically social network data.
2. Velocity: the data are created in real or near real time, e.g., e-commerce and back transactions.
3. Variety: the data are available in different types and shapes, such as structured, unstructured, and semi-structured data.

Smart cities are linked to technological advancements, with technologies such as big data, the Internet of Things (IoT), and machine learning enabling the development of smart cities and are interconnected. As the number of IoT devices and sensors is increasing dramatically, generating a huge volume of data, which is exponentially growing big data, machine learning becomes very necessary to help in collecting, processing, and making sense of the huge amount of data to gain insights and enable effective decision making (Moreno et al., 2017). However, some authors highlight the inadequacy of current machine learning solutions. Mohammadi et al. (2018), for instance, underline the problem of wasting unlabelled data, highlighting the importance of semi-supervised learning for smart cities to overcome this challenge.

Al Nuaimi et al. (2015) identified three main benefit areas of big data in smart cities, namely efficient resource utilization, improved quality of life, and higher level of openness and transparency. Allam & Dhunny (2019) emphasized that, on the one hand, artificial intelligence is needed to efficiently manage big data, but, on the other hand, human dimensions cannot be lost in smart cities and liveability dimensions should be supported by this technology. Hashem et al. (2016) introduced a vision of how big data can be analysed to support smart city objectives, relying on security and data science models. Lim et al. (2018) identified four reference models based on urban big data use cases, namely local operations management, preventive local administration, local information diffusion, and local network development, referring to different industry areas of big data applications. According to Bibri (2019), there are significant opportunities for big data analytics and its use in future smart cities to enhance their contribution to sustainable development goals. This can be achieved by optimising and improving urban services, functions, operations, policies and strategies, as well as addressing complex analytical questions and advancing knowledge domains. However, along with the vast opportunities to be embraced and harnessed, there are significant challenges that need to be addressed in order to successfully implement big data technology and its innovative applications in cities, including technical, analytical, computational, and logistic challenges. Using the value chain approach, Löfgren & Webster (2020) argue that many challenges are related to the technocratic perspective of governance, embedded norms, and emerging surveillance practices. The ethical requirements for the use of big data in smart cities were highlighted by (V. Chang, 2021), categorizing current ethical issues into (a) privacy and surveillance, (b) data integrity, (c) social equity and the aging population, and (d) public transportation.

Several studies have explored the use of big data in smart cities while identifying pertinent challenges in specific application areas such as healthcare (Pramanik et al., 2017), public safety (Liu et al., 2022) and transportation (Zhu et al., 2019). For example, Pramanik et al. (2017) presents a structured assessment of various technologies related to big data and intelligent systems, along with a thorough examination of advanced healthcare systems. From this analysis, three main technical branches (machine learning, intelligent agents, and text mining) are identified as significant contributors to the advancement of healthcare systems. In particular, a framework for a big data-enabled smart healthcare system is proposed, including theoretical models that represent intra- and inter-organisational business dynamics in the healthcare domain. Zhu et al. (2019) examined the attributes of big data and intelligent transport systems (ITSs), and present a framework for implementing big data analytics in ITSs. The framework outlines various categories of big data analytics applications. In addition, the authors present several case studies that demonstrate the application of big data analytics in ITSs, such as road traffic flow prediction and road accident analysis. Using a design science paradigm, Liu et al. (2022) proposed a model-driven approach to scenario modelling in government big data governance. This approach aims to enable flexible scenario modelling in the context of government operations.

Big data has diverse applications in different aspects of smart cities, including economy, governance, transportation, environment, and energy (Soomro et al., 2019). In terms of smart economy, big data enables the management and analysis of huge amounts of historical economic data, which contributes to the analysis and prediction of future economic scenarios, with economic forecasting being a key application. In smart governance, big data analytics enables the extraction of citizen preferences and opinions from multiple sources such as social media, facilitating data-driven strategies in city processes. Evidence-based decision making is facilitated by the integration and analysis of cross-sectoral data. In smart transport, big data analytics can help address traffic congestion by identifying alternative routes in advance, optimising traffic parking, and providing timely medical services after accidents. In the area of smart environments, big data analytics helps to collect, manage and analyse data from sources such as climate data and environmental sensors. Big data analytics also plays a crucial role in smart energy by promoting a cleaner environment, energy recycling, and reduced energy consumption.

Despite the growing interest in big data applications, few review articles have considered their use in smart cities (Table 1). To date, research on the use of big data in smart cities has tended to focus on systematic review of the literature. Chauhan et al. (2016) offer a comprehensive overview of key challenges associated with big data in smart cities, namely volume, variety, velocity, value, veracity, privacy, and human dynamics. Furthermore, a rigorous process applicable to effective big data management in smart cities is provided to big data practitioners. Karimi et al. (2021) introduced a taxonomy of algorithm- and architecture-based studies. Algorithm-based approaches prevailed, with meta-heuristic as the most widely used algorithm. Testbeds and simulation were identified as the most commonly used evaluation techniques. Osman (2019) reviewed 30 papers on big data analytics in smart cities, focusing on the functional requirements and value chain. The aim was to determine the characteristics of big data analytics frameworks used in smart cities, and to identify the design principles needed to serve their purposes. As a result, a new framework called the Smart City Data Analytics Panel was proposed based on the findings of the study, consisting of three layers: platform, security and data processing. Tiwari et al. (2019) performed a bibliometric analysis of a large number of articles related not only to big data but also to Industry 4.0 and IoT to introduce a conceptual framework for the effective deployment of Industry 4.0 technologies in smart cities. Unlike previous studies, this study uses content analysis of recent literature to identify the most influential research areas and analyse their evolution. This allows us not only to avoid subjective bias but also to reveal the most recent research trends and future research directions.

Table 1: Comparison of our study with previous literature reviews

Study	**Time period**	**Data**	**Motivation**	**Methodology**
Chauhan et al. (2016)	2012-2016	38 articles from multiple databases	identify challenges of leveraging big data in smart cities	systematic review
Allam & Dhunny (2019)	1986-2018	57 articles from multiple databases	present a comprehensive framework for incorporating big data and AI within smart cities, with the primary objective of enhancing livability	systematic review
Soomro et al. (2019)	2009-2018	66 articles from multiple databases	identify computing models, data models, privacy and security, and market drivers	technological and thematic analysis
Tiwari et al. (2019)	2013-2018	774 articles from Scopus database	explore the role of Industry 4.0 in smart cities	bibliometric analysis

Osman (2019)	2012-2017	30 articles from multiple databases	identify the features of big data analytics frameworks used within smart cities	systematic review
Karimi et al. (2021)	2013-2021	46 articles from Google Scholar	identify algorithm- and architecture-based approaches	systematic review
This study	2017-2022	60 articles from WoS and IEEE Xplore	identify the most influential research areas and their evolution	content analysis

3. Research Methodology

A literature review on the use of big data in smart cities was carried out to answer the following research questions (RQs):

RQ1: What research areas are explored in recent articles on the application of big data in smart cities?

RQ2: What is the evolution of research in the application of big data in smart cities?

To answer these research questions, we employed the following research methodology:

1. Research preparation (selection of the search strategy).
2. Data collection from databases and selection of articles included in content analysis.
3. Select words and phrases from articles in terms of frequency of occurrence.
4. Visualize science maps based on words and phrases using VOSviewer 1.6.9 (van Eck & Waltman, 2017).
5. Identify research areas and trends by visualizing their timelines in CiteSpace 6.1 (C. Chen et al., 2010).

First, a comprehensive search of two databases, the WoS Core Collection and the IEEE Xplore database, was conducted. To identify relevant articles, the search criteria were used according to Karimi et al. (2021) as follows: TS=(("big data" AND "smart" AND ("cities" OR "city")).
Articles were retrieved for the period January 2017 to June 2022 due to the increasing number of articles on big data in the last five years. The search resulted in 296 articles. After the screening process and eligibility check, 192 articles were included in the content analysis. Consistent with previous review studies (Hajek et al., 2022; Mora et al., 2019), we used content analysis to identify clusters of research areas using the cooccurrence of relevant terms. To further visualize the evolution of the research areas, a progressive network analysis was performed that considers the temporal aspects of the clusters produced. Thus, new research areas can be identified.

We then conducted a qualitative assessment to discard articles that were not relevant to the eight dimensions of smart cities. We reached a final number of 60 articles for the analysis of big data for smart city dimensions (Fig. 1).

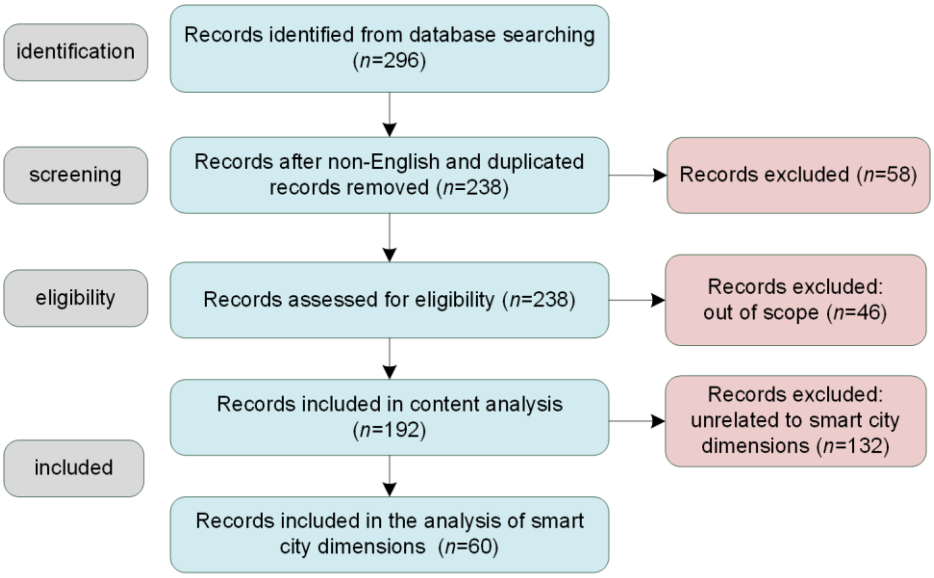

Fig. 1: Flow chart of the selection process

4 Results

4.1 Content Analysis

To answer *RQ1 (What research areas are explored in recent articles on the application of big data in smart cities?)*, cooccurrence analysis of terms was performed. The science maps were produced using terms (words and sequences of words - phrases) present in article titles and abstracts.

The analysis of the science map shown in Figure 2 resulted in five major clusters of terms listed in the title or abstract for all articles. The largest cluster in blue was related to big data, smart cities, and IoT, indicating a strong interconnection between the three terms. The second cluster in purple was focused on smart city and the mechanics of big data within the smart city with keywords such as system, network, and social media. The third cluster in red represents the Internet and IoT with related terms, such as things, information, cloud computing, and model. The second and third clusters emphasize a focus on data creation, generation, and management. The fourth cluster in yellow was related to big data analytics, innovation, services, and sustainability. The fifth and final cluster in green was the most decentralized, around keywords such as artificial intelligence, machine learning, algorithm, classification, prediction, security, and privacy. The fourth and fifth clusters highlight a focus on leveraging data through data analytics, machine learning, and artificial intelligence techniques.

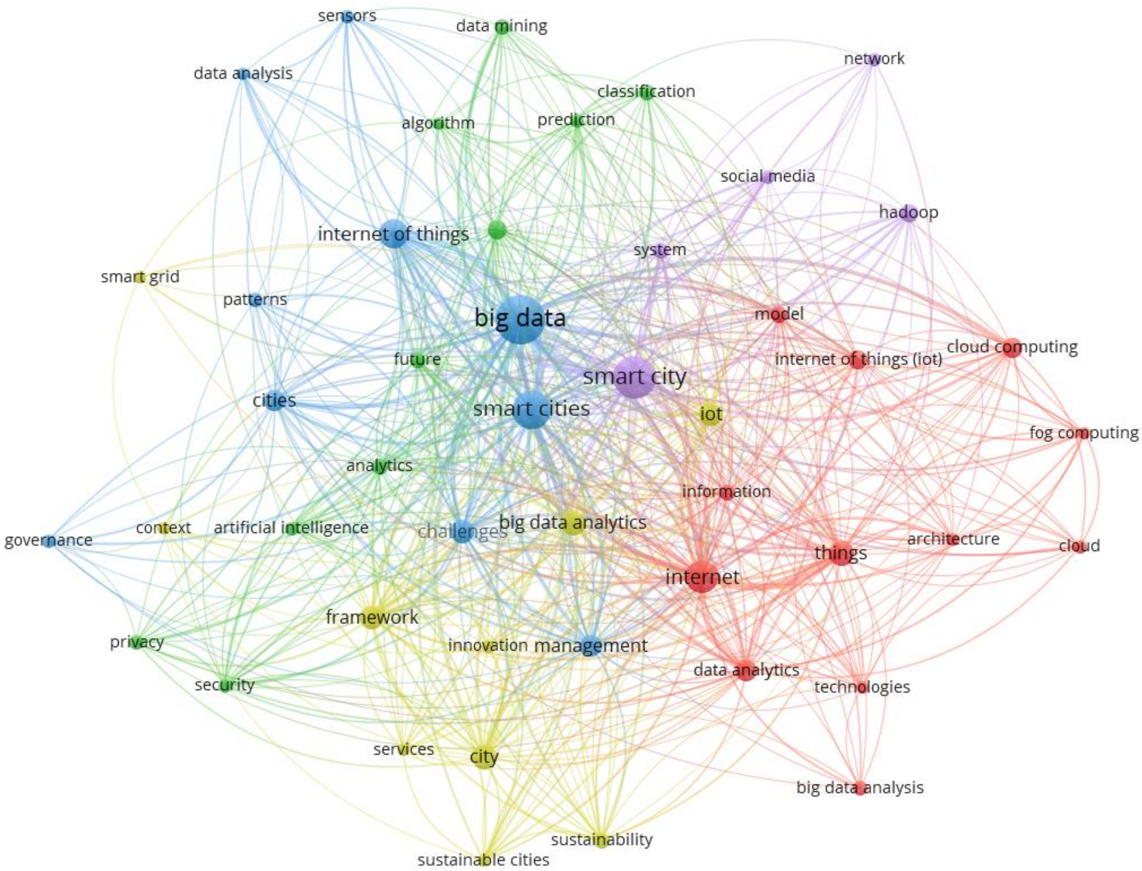

Fig. 2: Science map based on the co-occurrence of terms in titles and abstracts. A full count with a minimum number of occurrences of 5 was used. To identify terms (noun phrases), the Apache OpenNLP library was used to detect sentences and perform part-of-speech tagging. The map was constructed based on the similarity matrix represented by the association strengths between the occurrences of the terms.

To answer *RQ2 (What is the evolution of research in the application of big data in smart cities?),* progressive network analysis was conducted. Figure 3 indicates that seven clusters were produced on the science map. Each cluster represents a research area over the past five years. The evolution of each research area also highlights the relevant terms for years, in which the corresponding articles were highly cited.

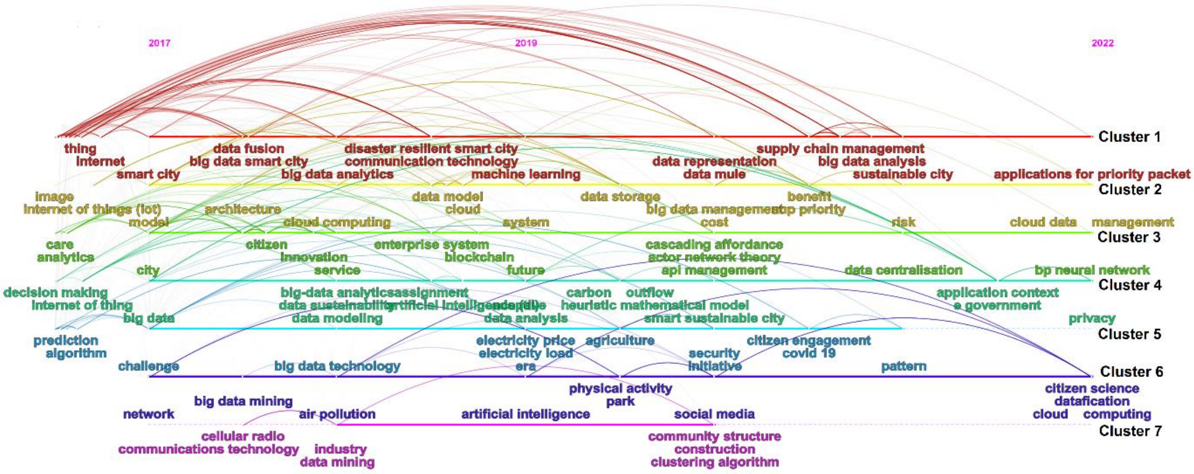

Fig. 3: Evolution of research areas. Terms were extracted from titles, keywords, and abstracts; nodes represent terms and articles; *g*-index was used to select nodes with a scale factor of 25 and the cosine link similarity. The merged network included 219 nodes and 905 edges (density = 0.038) with modularity Q = 0.519, weighted mean silhouette S = 0.864, and harmonic mean (Q, S) = 0.648.

Table 2 reports the results of the clustering of the science map. For each cluster, the number of articles denotes the size of the cluster, the silhouette value indicates the quality of the cluster, and the mean value of the year indicates when the cluster was formed. Regarding the quality of clustering, high silhouette values indicate good quality for all clusters. The overall quality of the clusters was S=0.865, calculated as the weighted mean silhouette index. To label the clusters, the log-likelihood ratio of terms obtained from article titles was used.

The research area of cluster 1 was "Using big data analytics", with 44 articles and the average year was 2019. However, as shown in Figure 3, this area of research draws on much earlier research on IoT and smart cities. In contrast, cluster 7 ("Wireless big data") emerged in 2018 and lasted until 2020. The most recently emerged research area was cluster 6 ("Green space"). On the one hand, the lack of strong vertical connections among the identified clusters in Figure 3 also suggests that the research areas are not well connected, even in terms of common research foundations. On the other hand, the relatively high value of betweenness centrality of cluster 4 suggests that the concept of smart sustainable cities is potentially critical to understanding the remaining clusters.

Table 2: Clusters obtained from the science map

Cluster	Size	Silhouette	Betweenness centrality*	Mean year	Label
1	44	0.896	0.23	2019	Using big data analytics
2	32	0.831	0.22	2019	Electric vehicle integration
3	31	0.794	0.19	2018	Citizen-centric big data
4	30	0.889	0.25	2019	Smart sustainable cities
5	28	0.811	0.16	2018	Deep learning
6	17	0.945	0.15	2020	Green space
7	10	0.988	0.10	2018	Wireless big data

* CiteSpace utilizes betweenness centrality in accordance with the principles of the structural hole theory.

4.2 Discussion on Identified Clusters

In the following, we discuss the identified areas of research with a focus on identifying potential future research trends in these areas. We will limit ourselves to six largest clusters from the progressive network analysis, as the last cluster does not seem to be an active research area anymore according to the results.

Cluster 1: Using big data analytics

During the past 5 years, using big data analytics in smart cities has attracted considerable attention from researchers. The articles in this cluster aim to explore the potential of using big data analytics in case studies of smart cities. For example, Xie et al. (2019) investigated how the links among big data, IoT, machine-to-machine, and smart cities can help support and meet the needs of retrofitting pilot projects with particular focus on traffic congestion management and smart grid. Based on these case studies, the application guidelines for the use of big data analytics in retrofitting projects are proposed for smart city decision makers. Other studies develop IoT-based systems using big data analytics to assist smart city planners. For example, Rathore et al. (2017) developed a four-layer architecture to handle data collection and generation (bottom tier), communication (intermediate tier-1), data administration and processing (intermediate tier-2), and data interpretation (top tier). This system architecture was later applied to smart transportation, using data from road sensor and vehicular network, building and processing graphs, and providing aggregated interpretable results for decision making in several domains, such as traffic and accident management, over-speed analysis, travellers guidance, and crowded road search (Rathore et al., 2018). Finally, several challenges of using big data analytics in smart cities were identified by Tiwari et al. (2019), namely (1) data quality (trustworthy sources of data should be identified), (2) data integration (technical expertise is required to integrate data from different sources), (3) data privacy (an essential aspect for citizens), (4) understanding citizens' needs (valued information should be determined), (5) geographic modification (geographic information is often missing), and (6) service design (the quality of urban services should be improved).

Cluster 2: Electric vehicle integration

Electric vehicles produce and use big data, helping smart cities become green by reducing emissions (B. Li et al., 2017). Big data analytics was reported to be effective in integrating electric vehicles in several ways, including battery management, optimized charging, vehicle status tracking (B. Li et al., 2017). Big data were also used to optimize electric vehicle transportation networks, which significantly reduced congestion spread (Lv et al., 2021). In other studies, electric mobility as a service (eMaaS) is emphasized as a way to make transportation more effective and greener (Anthony Jnr et al., 2020). Big data architectures can then be utilized to collect, process, analyse and use mobility data for eMaaS in smart cities. However, most previous big data technologies in this area relied on Hadoop clusters in the cloud, which is a concern for processing data in real time. Indeed, due to the unprecedented increase in the number of IoT devices, increasing cloud computing servers face the issues of privacy, response time, storage, and security. These issues can be addressed thanks to the introduction of emerging paradigms such as edge and fog computing, which embrace the concept of IoT devices (Badidi et al., 2020). This is particularly relevant for smart mobility (Pereira et al., 2019).

Cluster 3: Citizen-centric big data

A citizen-centred big data framework was introduced by (Ju et al., 2018), consisting of two diverse perspectives, data analysis algorithms and urban governance. In the first layer of the framework, big data are merged to create a citizen-centric panoramic dataset. Then, individual and group citizen profiles and personas are designed, and data analysis algorithms are applied to discover knowledge in the data. In the final layer, knowledge models are produced and optimization algorithms are used to assist decision makers via an urban governance intelligence ontology model. It is argued that citizen-centred data can provide a better understanding of public opinion and citizen behaviour (Ju et al., 2018). In this line of research, Jun et al. (2018) performed a big data analysis for the smart city of Namyangju, Korea, with the intention of increasing the level of satisfaction with civil services that respond to the needs and demands of newly arrived residents. Hence, it is demonstrated how big data implementation can support smart city policy and, in turn, improve the quality of life in a city. Similarly, Kim et al. (2021)

present a text mining approach to analyse online civic query data. Thus, civic complaints can be automatically identified in the large number of 160,000 civic queries, and the quality of governance response can be improved.

Cluster 4: Smart sustainable cities

One of the most demanding issues in the current trend of smart cities is keeping environmental aspects under control with the ultimate goal of sustainability. Therefore, the concept of smart sustainable cities was introduced (Huovila et al., 2019), and big data technologies were exploited in smart city infrastructures to support environmental monitoring (Garau et al., 2020). In addition, predictive models can be used to control pollution levels to protect citizens from the damage caused by pollution (Honarvar & Sami, 2019). Such pollution predictions can be utilized in smart mobility and traffic management (J. Chang et al., 2020). Khan et al. (2022) identified some major barriers to big data that need to be addressed in the development of smart sustainable cities, such as the lack of technologies and frameworks for big data analytics.

Cluster 5: Deep learning

IoT big data processing is an increasingly growing industry, and many researchers are passionate about trying to make cities smarter by utilizing modern machine learning methods. Among these methods, deep learning-based methods dominate over existing machine learning methods, which have recently shown compelling results, especially for multimedia data (Bellavista et al., 2020). The applications of deep learning in smart cities include the design of distributed cyberinfrastructure to integrate city-level social networks and crowd-sourced data to improve public safety and quality of life (Shams et al., 2018). Distributed deep learning-based architectures for big data systems in smart cities were also used for toluene prediction (smart environment) and wind energy prediction (smart energy) (Naoui et al., 2021). Atitallah et al. (2020) reviewed a taxonomy of deep learning models and tools for IoT data, showing their pros and cons along with commonly used deep learning datasets in smart city tasks.

Cluster 6: Green space

Urban green space is considered a major component of the smart environment, providing ecosystem services such as tourism and recreation, agriculture, and air purification (Dwevedi et al., 2018). The design and planning of green spaces require a variety of social, ecological, and biological data. Chinese microblog data were used to model the temporal and spatial behaviour of park visitors so that different behavioural patterns and green space categories can be identified (Q. Liu, Ullah, Wan, Peng, Hou, Rizvi, et al., 2020). Similarly, to avoid time-consuming surveys, location-based social network data were utilized to understand spatiotemporal behaviour of green park visitors in Shanghai, underlining the role of social media data in urban planning in agreement with citizen preferences (Q. Liu, Ullah, Wan, Peng, Hou, Qu, et al., 2020).

4.3 Smart City Dimensions

As shown in Table 3, the 60 articles we selected for further analysis were divided into 8 dimensions: smart mobility, smart environment, smart energy, smart infrastructure, smart healthcare, smart governance, smart people and smart economy. In the following, we discuss the impact of big data on each dimension.

Table 3: Summary of the reviewed articles based on the dimensions being addressed. Manual categorization of the articles was used to provide a more nuanced and context-specific categorization in terms of smart city dimensions.

Dimension of smart city	No. of articles	References
Smart mobility	20	(Abberley et al., 2017; Schatzinger & Lim, 2017; Sharif et al., 2017; Chen et al., 2018; X. Cui et al., 2019; Leung et al., 2019; Murk et al., 2019; Hussain et al., 2020; Kanwal et al., 2020; Mystakidis & Tjortjis, 2020; Asaithambi et al., 2020; Shin, 2020; Zhou et al., 2020; W. Li & Chen, 2021; Feng et al., 2021; Lv et al., 2021; Cavicchioli et al., 2022; Aljehane & Mansour, 2022)
Smart environment	13	(B. Li et al., 2017; Navarro et al., 2017; Bibri, 2018; Zaree & Honarvar, 2018; Aifadopoulou et al., 2019; Honarvar & Sami, 2019, 2021; Garau et al., 2020; Lavalle et al., 2020; Q. Liu, Ullah, Wan, Peng, Hou, Qu, et al., 2020; Z. Yang et al., 2021; Kaginalkar et al., 2022)
Smart energy	10	(Tian et al., 2017; B. Li et al., 2017; Qiu et al., 2017; Mohamed et al., 2018; Nandury & Begum, 2018; Pérez-Chacón et al., 2018; Mujeeb et al., 2019; Murk et al., 2019; Sayah et al., 2020; Arif et al., 2022)
Smart infrastructure	8	(C. Yang et al., 2017; Rani & Chauhdary, 2018; Wei & Yang, 2018; T. Wang et al., 2018; Silva et al., 2020; W. Wang, 2020; Zhang et al., 2021; Sharma et al., 2022)
Smart healthcare	7	(Islam et al., 2017; Obinikpo & Kantarci, 2017; Pramanik et al., 2017; Anisetti et al., 2018; Clim et al., 2019; EL Azzaoui et al., 2021; Kamel Boulos & Koh, 2021)
Smart governance	5	(C. Yang et al., 2017; Kourtit & Nijkamp, 2018; Caputo et al., 2019; Nica et al., 2020; Haidery et al., 2020)
Smart people	2	(Dooley, 2021; Kim et al., 2021)
Smart economy	1	(Xiao & Xie, 2021)

Smart mobility

Today, vehicles are connected to the Internet in some way or another and generate a lot of data. These data can be processed to provide a wealth of information, such as vehicle type, traffic levels, and congestion, which can help citizens to better anticipate and plan their journeys and ease mobility problems. The number of electric vehicles is also growing, which is one of the elements of smart cities. Electric vehicles can help generate different types of data that lead to improvements not only in mobility but also in other sectors such as energy by optimising charging methods and the environment by providing information on emission levels and working to reduce them. We identified 20 articles that address the use of big data in the context of smart mobility, including road congestion modelling (Abberley et al., 2017), vehicle routing (Aljehane & Mansour, 2022), autonomous vehicle management (Cui et al., 2019), shared electric vehicle use (Feng et al., 2021), electric vehicle integration (Hussain et al., 2020), smart parking systems (W. Li & Chen, 2021), and smart traffic management (Sharif et al., 2017).

Smart environment

The environment and sustainability have always been key drivers of smart city development (Akande et al., 2019). Smart cities strive to protect the environment and conserve natural resources. The IoT has specific use cases to enable environmentally friendly and sustainable production and consumption patterns under the so-called Green Internet of Things, "G-IoT". G-IoT can use data to help monitor pollution levels, resource consumption, and share resources, with better insight enabling citizens and government authorities to make better environmental decisions. The identified articles in this dimension mainly present analytical networks for IoT-based big data applications (Bibri, 2018; Garau et al., 2020; Navarro et al., 2017).

Smart energy

Energy fuels many infrastructures and resources. Cities have a growing demand for energy, which drives increased production and development. Smart cities are working to use big data analytics to increase the level of efficiency in energy production and consumption. These efficiencies can be applied to monitoring and controlling energy networks through several thousand sensors. The articles classified in this dimension comprise applications such as automatic detection of electricity theft (Arif et al., 2022), smart grid operation (Nandury & Begum, 2018), energy management of buildings (Mohamed et al., 2018), electricity price and load forecasting (Mujeeb et al., 2019), and the analysis of behavioural patterns of electricity consumption (Pérez-Chacón et al., 2018).

Smart infrastructure

Infrastructure is a very broad dimension that can encompass many different sectors. If focused on some physical infrastructure such as buildings, roads, and bridges, a large amount of data can be collected and used to help smart cities monitor the quality of infrastructure and proactively predict and prevent potential damage. Articles in this domain deal with smart infrastructure oriented to IoT (Sharma et al., 2022), the Web of Things (Silva et al., 2020), infrastructural health monitoring (T. Wang et al., 2018), and enhancement of critical infrastructure in order to promote smart city resilience (C. Yang et al., 2017; Zhang et al., 2021).

Smart healthcare

Healthcare services are very important to people in terms of their quality of life and their satisfaction with the city in which they live. Big data can drive smarter healthcare systems, as data can be generated from different devices connected to people and then fed into healthcare systems, allowing real-time monitoring and effective decision-making on patient health. Several applications based on big data collection and ICT technologies such as mobile and cloud technologies have been developed in this area. Privacy-aware big data analytics can serve as a valuable tool supporting public health policies (Anisetti et al., 2018; Clim et al., 2019), including prediction of Covid-19 outbreaks (EL Azzaoui et al., 2021).

Smart governance

City management is constantly looking for ways to make effective, quick, and informed decisions. Big data can play a role here and support the decision-making process through solutions such as an interactive dashboard that reveals more and faster information about the performance of different sectors, informs different stakeholders, highlights patterns, highlights issues that need to be fixed, and suggests corrective actions. The identified articles handle the following issues: residential density analysis (Haidery et al., 2020), introduction of smart city ecosystems (Caputo et al., 2019), decision-support systems using big data dashboards (Kourtit & Nijkamp, 2018), and crisis management (C. Yang et al., 2017). On the one hand, technology can have a positive effect on the balance of power. Technology can facilitate collaboration and co-creation between citizens, mayors and city administrators, and online platforms enable collective problem-solving and policy-making collaboration. On the other hand, ensuring equal access to technology and digital literacy are essential to prevent disadvantaged communities from being further marginalized. In addition, to protect citizens' rights and maintain trust in technology-enabled governance, strong data protection and privacy measures should be in place.

Smart people

The modern era of smart cities is citizen-centric. Active citizens and their participation are therefore crucial to the development of a smart city. Governments that manage smart cities are developing many electronic and on-line public services to facilitate citizen engagement and provide quality services at their fingertips. Seoul is one of the cities where a large amount of data is collected from citizens through online platforms. These data represent citizen interactions with e-government services, including requests, suggestions, and complaints. The volume of data can be hundreds of thousands or millions of records over the years, where it can be analysed and used to gain insight from the citizens' perspective on how the city can adjust its urban planning and policies to better meet the needs and desires of its citizens (Kim et al., 2021). From the perspective of passive participation, citizens are viewed as sensors and sources for big data analytics (Dooley, 2021).

Smart economy

Economic development is the result of the development and management of all other dimensions. In other words, smart governance will lead to a more efficient and higher use of city resources and consequently economic growth; similarly, a healthy and intelligent population will lead to a more engaged, efficient, and productive population and, therefore economic growth. Several studies have looked at the impact of big data and smart city construction on economic growth, finding economic and social benefits (Zhao & Zhang, 2020). Indeed, the identified study found that smart cities and big data promote economic growth and improve the quality of urban development (Xiao & Xie, 2021).

Additional considerations

The eight dimensions above are indicative, some studies branch around them or regroup them in a different way to focus on a particular dimension. Therefore, smart technology can be considered as a dimension, a sub-dimension of smart infrastructure or an overall enabler. There is no doubt that the impact of technology is fundamental to the development and evolution of smart cities. However, making it the core part can lead to the risk of slowing down the development of smart cities or undermining the impact of the environment, sustainability or citizens' needs (Dana et al., 2022). Specifically, an over-emphasis on technology without considering broader issues of viability and sustainability can pose threats to cities, such as a widening of the 'digital divide', negative environmental impacts, a decline in community engagement and the creation of vulnerabilities in urban systems.

Ultimately, smart cities depend on the engagement of citizens and their acceptance of new concepts, while seeing value in responding to their needs. Here, culture plays a key role in citizen participation and acceptance of the whole ecosystem (Allam and Dhunny, 2019). In other words, culture influences the way people interact with their surroundings, their environment, including technology. Their behaviour, values, preferences and attitudes can accelerate or slow down certain dimensions of smart city development. And as this paper focuses on the role of big data in the development of smart cities, concepts such as privacy and digital trust are essential to harness the collection, storage, processing, and sharing of data, which in turn goes back to culture. For example, GDPR applies to all companies processing data of EU citizens, but may present some challenges for GAFA and its involvement in smart city development in the EU, including obtaining explicit consent, implementing privacy-by-design principles, and addressing data purpose and data location restrictions.

Similarly, smart healthcare is sometimes considered together with public safety as a key focus area of smart living (Allam and Dhunny, 2019). Our results suggest that the identified articles focus on smart healthcare rather than public safety, so we opted for the smart healthcare dimension and included public safety in the smart infrastructure dimension.

There are a number of factors that can make the application of big data in smart cities different between small, medium-sized, and large cities (Borsekova et al., 2018; Duygan et al., 2022). The infrastructure and the volume of data generated are typically more extensive in large cities. This scale provides the opportunity for a more comprehensive collection and analysis of data. Large cities, with their complex transport networks, diverse economic activities and varied social dynamics, tend to generate a more diverse range of data. As a result, more comprehensive insights into urban challenges and more targeted interventions can be gained from this diversity of data sources. However, even in smaller cities, the data that is available can still provide valuable insights for the improvement of services and decision making. Taking into account their unique context and challenges, they can adapt and customise technologies to their specific needs.

4.4 Future Research Directions

Taking into account the identified research areas and the dimensions of smart city discussed above, Table 4 presents the suggested future research directions for the identified research areas and the dimensions of smart cities.

Table 4: Suggested future research directions

Research area	**Future research directions**	**Smart city dimensions**
Using big data analytics	− Evaluate the economic impact of using big data analytics	smart economy
	− Integrate technology-centered and human-centered scenarios	smart people
Electric vehicle integration	− Evaluate the impact of edge / fog / edge-fog computing on smart energy and smart mobility	smart mobility, smart energy
	− Green fog computing in smart cities	smart infrastructure

Citizen-centric big data	− Deploy responsive governance in real time to meet the citizen needs	smart governance
	− Utilize state-of-the-art language and image models to understand citizen needs	smart people
Smart sustainable cities	− Develop big data analytics technologies and frameworks for smart sustainable cities	smart infrastructure
	− Exploit the nature of big data in developing smart sustainable cities	smart environment
Deep learning	− Develop distributed parallelism strategies for deep learning	smart infrastructure
	− Use computationally effective deep learning architectures for multimedia and multimodal data	smart infrastructure
Green space	− More research on the combination and visualization of demographic, biological, and environmental data from urban green space	smart environment
	− Integration of IoT into natural environment	smart environment

5. Conclusion

We can conclude that the use of big data can bring about a number of benefits for the development of smart cities. As mentioned in the results section, most existing studies have investigated the impact of big data on smart cities with use cases in different sectors, mostly trying to apply data analysis techniques to enable more efficient resource management and decision making.

Current research areas on the use of big data in smart cities were mainly in IoT, big data analytics, predictive analytics, and the challenges of big data in smart cities (RQ1). Progressive network analysis revealed several research areas that dominate the recent research in this field, especially using big data analytics, smart mobility, citizen-centric big data, smart sustainable cities, deep learning, and green space applications (RQ2). As a result of our evaluation, we mapped the 60 shortlisted studies that address the use of big data in smart cities onto eight smart city dimensions. It seems that the biggest focus in the use of big data in smart cities is on smart mobility with use cases related to traffic management, smart parking, and efficient transportation systems. Smart environment, energy, and infrastructure also seem to be attracting a lot of attention with several use cases related to environmental and sustainability monitoring, automated energy production and consumption management systems, as well as infrastructure monitoring. Several studies have addressed multiple dimensions, for example (Aifadopoulou et al., 2019). However, we expected to find more studies related to dimensions of smart economy and smart people, as we could identify only two studies related to smart people and only one study related to smart economy. This leads us to identify these two dimensions as the main research gaps and to suggest that future research should focus on areas related to value stream creation for the smart economy sector in smart cities based on big data such as the decentralized economy.

References

Abberley, L., Gould, N., Crockett, K., & Cheng, J. (2017). Modelling road congestion using ontologies for big data analytics in smart cities. *2017 International Smart Cities Conference, ISC2 2017*, 1–6. https://doi.org/10.1109/ISC2.2017.8090795

Aifadopoulou, G., Salanova, J. M., Tzenos, P., Stamos, I., & Mitsakis, E. (2019). Big and open data supporting sustainable mobility in smart cities – The case of Thessaloniki. *Advances in Intelligent Systems and Computing*, *879*, 386–393. https://doi.org/10.1007/978-3-030-02305-8_47/COVER/

Akande, A., Cabral, P., Gomes, P., & Casteleyn, S. (2019). The Lisbon ranking for smart sustainable cities in Europe. *Sustainable Cities and Society*, *44*, 475–487. https://doi.org/10.1016/J.SCS.2018.10.009

Al Nuaimi, E., Al Neyadi, H., Mohamed, N., & Al-Jaroodi, J. (2015). Applications of big data to smart cities. *Journal of Internet Services and Applications*, *6*(1), 1–15. https://doi.org/10.1186/S13174-015-0041-5/FIGURES/3

Aljehane, N. O., & Mansour, R. F. (2022). Big data analytics with oppositional moth flame optimization based vehicular routing protocol for future smart cities. *Expert Systems*, *39*(5), e12718. https://doi.org/10.1111/EXSY.12718

Allam, Z., & Dhunny, Z. A. (2019). On big data, artificial intelligence and smart cities. *Cities*, *89*, 80–91. https://doi.org/10.1016/J.CITIES.2019.01.032

Anisetti, M., Ardagna, C., Bellandi, V., Cremonini, M., Frati, F., & Damiani, E. (2018). Privacy-aware Big Data Analytics as a service for public health policies in smart cities. *Sustainable Cities and Society*, *39*, 68–77. https://doi.org/10.1016/J.SCS.2017.12.019

Anthony Jnr, B., Abbas Petersen, S., Ahlers, D., & Krogstie, J. (2020). Big data driven multi-tier architecture for electric mobility as a service in smart cities: A design science approach. *International Journal of Energy Sector Management*, *14*(5), 1023–1047. https://doi.org/10.1108/IJESM-08-2019-0001

Arif, A., Alghamdi, T. A., Khan, Z. A., & Javaid, N. (2022). Towards efficient energy utilization using big data analytics in smart cities for electricity theft detection. *Big Data Research*, *27*, 100285. https://doi.org/10.1016/J.BDR.2021.100285

Asaithambi, S. P. R., Venkatraman, R., & Venkatraman, S. (2020). MOBDA: Microservice-oriented big data architecture for smart city transport systems. *Big Data and Cognitive Computing 2020, Vol. 4, Page 17*, *4*(3), 17. https://doi.org/10.3390/BDCC4030017

Atitallah, S. Ben, Driss, M., Boulila, W., & Ghezala, H. Ben. (2020). Leveraging deep learning and IoT big data analytics to support the smart cities development: Review and future directions. *Computer Science Review*, *38*, 100303. https://doi.org/10.1016/J.COSREV.2020.100303

Badidi, E., Mahrez, Z., & Sabir, E. (2020). Fog computing for smart cities' big data management and analytics: A review. *Future Internet*, *12*(11), 1–29.

https://doi.org/10.3390/fi12110190

Batty, M. (2019). Urban analytics defined. *Environment and Planning B: Urban Analytics and City Science*, *46*(3), 403–405. https://doi.org/10.1177/2399808319839494

Bellavista, P., Ota, K., Lv, Z., Mehmood, I., & Rho, S. (2020). Towards smarter cities: Learning from Internet of Multimedia Things-generated big data. *Future Generation Computer Systems*, *108*, 879–881. https://doi.org/10.1016/J.FUTURE.2019.06.003

Bibri, S. E. (2018). The IoT for smart sustainable cities of the future: An analytical framework for sensor-based big data applications for environmental sustainability. *Sustainable Cities and Society*, *38*, 230–253. https://doi.org/10.1016/J.SCS.2017.12.034

Bibri, S. E. (2019). On the sustainability of smart and smarter cities in the era of big data: an interdisciplinary and transdisciplinary literature review. *Journal of Big Data*, *6*(1), 1–64. https://doi.org/10.1186/S40537-019-0182-7/TABLES/3

Borsekova, K., Koróny, S., Vaňová, A., & Vitálišová, K. (2018). Functionality between the size and indicators of smart cities: A research challenge with policy implications. *Cities*, *78*, 17–26. https://doi.org/10.1016/J.CITIES.2018.03.010

Caputo, F., Walletzky, L., & Štepánek, P. (2019). Towards a systems thinking based view for the governance of a smart city's ecosystem: A bridge to link Smart Technologies and Big Data. *Kybernetes*, *48*(1), 108–123. https://doi.org/10.1108/K-07-2017-0274/FULL/XML

Cavicchioli, R., Martoglia, R., & Verucchi, M. (2022). A novel real-time edge-cloud big data management and analytics framework for smart cities. *Journal of Universal Computer Science*, *28*(1), 3–26. https://doi.org/10.3897/JUCS.71645

Chang, J., Kadry, S. N., & Krishnamoorthy, S. (2020). Review and synthesis of Big Data analytics and computing for smart sustainable cities. *IET Intelligent Transport Systems*, *14*(11), 1363–1370. https://doi.org/10.1049/IET-ITS.2020.0006

Chang, V. (2021). An ethical framework for big data and smart cities. *Technological Forecasting and Social Change*, *165*, 120559. https://doi.org/10.1016/J.TECHFORE.2020.120559

Chauhan, S., Agarwal, N., & Kar, A. K. (2016). Addressing big data challenges in smart cities: a systematic literature review. *Info*, *18*(4), 73–90. https://doi.org/10.1108/INFO-03-2016-0012/FULL/XML

Chen, C., Ibekwe-SanJuan, F., & Hou, J. (2010). The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis. *Journal of the American Society for Information Science and Technology*, *61*(7), 1386–1409. https://doi.org/10.1002/ASI.21309

Chen, X., Wang, H. H., & Tian, B. (2018). Visualization model of big data based on self-organizing feature map neural network and graphic theory for smart cities. *Cluster Computing*, *22*(6), 13293–13305. https://doi.org/10.1007/S10586-018-1848-1

Clim, A., Zota, R., Constantinescu, R., & Ilie-Nemedi, I. (2019). Health services in smart cities: Choosing the big data mining based decision support. *International Journal of Healthcare Management*, *13*(1), 79–87. https://doi.org/10.1080/20479700.2019.1650478

Cui, Q., Wang, Y., Chen, K. C., Ni, W., Lin, I. C., Tao, X., & Zhang, P. (2019). Big data analytics and network calculus enabling intelligent management of autonomous vehicles in a smart city. *IEEE Internet of Things Journal*, *6*(2), 2021–2034. https://doi.org/10.1109/JIOT.2018.2872442

Dana, L.-P., Salamzadeh, A., Hadizadeh, M., Heydari, G., & Shamsoddin, S. (2022). Urban entrepreneurship and sustainable businesses in smart cities: Exploring the role of digital technologies. *Sustainable Technology and Entrepreneurship*, *1*(2), 100016. https://doi.org/10.1016/J.STAE.2022.100016

Dooley, K. (2021). Direct passive participation: Aiming for accuracy and citizen safety in the era of big data and the smart city. *Smart Cities*, *4*(1), 336–348. https://doi.org/10.3390/SMARTCITIES4010020

Duygan, M., Fischer, M., Pärli, R., & Ingold, K. (2022). Where do smart cities grow? The spatial and socio-economic configurations of smart city development. *Sustainable Cities and Society*, *77*, 103578. https://doi.org/10.1016/J.SCS.2021.103578

Dwevedi, R., Krishna, V., & Kumar, A. (2018). Environment and big data: Role in smart cities of India. *Resources*, *7*(4), 64. https://doi.org/10.3390/RESOURCES7040064

EL Azzaoui, A., Singh, S. K., & Park, J. H. (2021). SNS big data analysis framework for COVID-19 outbreak prediction in smart healthy city. *Sustainable Cities and Society*, *71*, 102993. https://doi.org/10.1016/J.SCS.2021.102993

Feng, Y., Zhong, D., Sun, P., Zheng, W., Cao, Q., Luo, X., & Lu, Z. (2021). Micromobility in smart cities: A closer look at shared dockless e-scooters via big social data. *IEEE International Conference on Communications*, 1–6. https://doi.org/10.1109/ICC42927.2021.9500821

Garau, C., Nesi, P., Paoli, I., Paolucci, M., & Zamperlin, P. (2020). A big data platform for smart and sustainable cities: Environmental monitoring case studies in Europe. *Lecture Notes in Computer Science*, *12255*, 393–406. https://doi.org/10.1007/978-3-030-58820-5_30/COVER/

Haidery, S. A., Ullah, H., Ullah Khan, N., Fatima, K., Shahla Rizvi, S., & Kwon, S. J. (2020). Role of big data in the development of smart city by analyzing the density of residents in Shanghai. *Electronics*, *9*(5), 837. https://doi.org/10.3390/ELECTRONICS9050837

Hajek, P., Youssef, A., & Hajkova, V. (2022). Recent developments in smart city assessment: A bibliometric and content analysis-based literature review. *Cities*, *126*, 103709. https://doi.org/10.1016/J.CITIES.2022.103709

Hashem, I. A. T., Chang, V., Anuar, N. B., Adewole, K., Yaqoob, I., Gani, A., Ahmed, E., & Chiroma, H. (2016). The role of big data in smart city. *International Journal of Information Management*, *36*(5), 748–758. https://doi.org/10.1016/J.IJINFOMGT.2016.05.002

Honarvar, A. R., & Sami, A. (2019). Towards sustainable smart city by particulate matter prediction using urban big data, excluding expensive air pollution infrastructures. *Big Data Research*, *17*, 56–65. https://doi.org/10.1016/J.BDR.2018.05.006

Honarvar, A. R., & Sami, A. (2021). Particular matter prediction using synergy of multiple source urban big data in smart cities. *Intelligent Decision Technologies*, *15*(3), 371–385. https://doi.org/10.3233/IDT-200147

Huovila, A., Bosch, P., & Airaksinen, M. (2019). Comparative analysis of standardized indicators for Smart sustainable cities: What indicators and standards to use and when? *Cities*, *89*, 141–153. https://doi.org/10.1016/J.CITIES.2019.01.029

Hussain, M. M., Beg, M. M. S., Alam, M. S., & Laskar, S. H. (2020). Big data analytics platforms for electric vehicle integration in transport oriented smart cities: Computing platforms for platforms for electric vehicle integration in smart cities. In *Cyber Warfare and Terrorism: Concepts, Methodologies, Tools, and Applications* (pp. 833–854). IGI Global. https://doi.org/10.4018/978-1-7998-2466-4.CH051

Islam, M. M., Razzaque, M. A., Hassan, M. M., Ismail, W. N., & Song, B. (2017). Mobile cloud-based big healthcare data processing in smart cities. *IEEE Access*, *5*, 11887–11899. https://doi.org/10.1109/ACCESS.2017.2707439

Jorgensen, A., Rowland-Jones, J., Welch, J., Clark, D., Price, C., & Mitchell, B. (2014). *Microsoft Big Data Solutions* (1st ed.). John Wiley & Sons. https://books.google.cz/books?hl=cs&lr=&id=ISLwAgAAQBAJ&oi=fnd&pg=PT14&dq=Microsoft+Big+Data+Solutions&ots=6mgL2jiRiD&sig=AH215E_f_nirybYqFaowfFa9D40&redir_esc=y#v=onepage&q=Microsoft Big Data Solutions&f=false

Ju, J., Liu, L., & Feng, Y. (2018). Citizen-centered big data analysis-driven governance intelligence framework for smart cities. *Telecommunications Policy*, *42*(10), 881–896. https://doi.org/10.1016/J.TELPOL.2018.01.003

Jun, D., Hagen, L., Lee, E., Lim, H., & Kim, D. (2018). How big data contributes to the building of citizen-centric smart cities: The case of Namyangju City in Korea. *ACM International Conference Proceeding Series*, 1–6. https://doi.org/10.1145/3209281.3209358

Kaginalkar, A., Kumar, S., Gargava, P., Kharkar, N., & Niyogi, D. (2022). SmartAirQ: A big data governance framework for urban air quality management in smart cities. *Frontiers in Environmental Science*, *10*, 228. https://doi.org/10.3389/FENVS.2022.785129/BIBTEX

Kamel Boulos, M. N., & Koh, K. (2021). Smart city lifestyle sensing, big data, geo-analytics and intelligence for smarter public health decision-making in overweight, obesity and type 2 diabetes prevention: the research we should be doing. *International Journal of Health Geographics*, *20*(1), 1–10. https://doi.org/10.1186/S12942-021-00266-0/FIGURES/1

Kandt, J., & Batty, M. (2021). Smart cities, big data and urban policy: Towards urban analytics for the long run. *Cities*, *109*, 102992. https://doi.org/10.1016/J.CITIES.2020.102992

Kanwal, M., Malik, A. W., Rahman, A. U., Mahmood, I., & Shahzad, M. (2020). Sustainable vehicle-assisted edge computing for big data migration in smart cities. *IEEE Internet of Things Journal*, *7*(3), 1857–1871. https://doi.org/10.1109/JIOT.2019.2957127

Karimi, Y., Haghi Kashani, M., Akbari, M., & Mahdipour, E. (2021). Leveraging big data in smart cities: A systematic review. *Concurrency and Computation: Practice and Experience*, *33*(21), e6379. https://doi.org/10.1002/CPE.6379

Khan, M. A., Siddiqui, M. S., Rahmani, M. K. I., & Husain, S. (2022). Investigation of big data analytics for sustainable smart city development: An emerging country. *IEEE Access*, *10*, 16028–16036. https://doi.org/10.1109/ACCESS.2021.3115987

Kim, B., Yoo, M., Park, K. C., Lee, K. R., & Kim, J. H. (2021). A value of civic voices for smart city: A big data analysis of civic queries posed by Seoul citizens. *Cities*, *108*, 102941. https://doi.org/10.1016/J.CITIES.2020.102941

Kourtit, K., & Nijkamp, P. (2018). Big data dashboards as smart decision support tools for i-cities – An experiment on stockholm. *Land Use Policy*, *71*, 24–35. https://doi.org/10.1016/J.LANDUSEPOL.2017.10.019

Kourtit, K., Nijkamp, P., & Steenbruggen, J. (2017). The significance of digital data systems for smart city policy. *Socio-Economic Planning Sciences*, *58*, 13–21. https://doi.org/10.1016/J.SEPS.2016.10.001

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, *949*, 1–4.

Lavalle, A., Teruel, M. A., Maté, A., & Trujillo, J. (2020). Improving Sustainability of Smart Cities through Visualization Techniques for Big Data from IoT Devices. *Sustainability*, *12*(14), 5595. https://doi.org/10.3390/SU12145595

Leung, C. K., Braun, P., Hoi, C. S. H., Souza, J., & Cuzzocrea, A. (2019). Urban analytics of big transportation data for supporting smart cities. *Lecture Notes in Computer Science*, *11708*, 24–33. https://doi.org/10.1007/978-3-030-27520-4_3/COVER/

Li, B., Kisacikoglu, M. C., Liu, C., Singh, N., & Erol-Kantarci, M. (2017). Big data analytics for electric vehicle integration in green smart cities. *IEEE Communications Magazine*, *55*(11), 19–25. https://doi.org/10.1109/MCOM.2017.1700133

Li, W., & Chen, L. (2021). Habits in the transformation of smart parking system in the old city under the background of big data + humanistic psychology. *Psychiatria Danubina*, *33*, 18–569.

Lim, C., Kim, K. J., & Maglio, P. P. (2018). Smart cities with big data: Reference models, challenges, and considerations. *Cities*, *82*, 86–99. https://doi.org/10.1016/J.CITIES.2018.04.011

Liu, Q., Ullah, H., Wan, W., Peng, Z., Hou, L., Qu, T., & Haidery, S. A. (2020). Analysis of green spaces by utilizing big data to support smart cities and environment: A case study about the city center of Shanghai. *ISPRS International Journal of Geo-Information*, *9*(6), 360. https://doi.org/10.3390/IJGI9060360

Liu, Q., Ullah, H., Wan, W., Peng, Z., Hou, L., Rizvi, S. S., Haidery, S. A., Qu, T., & Muzahid, A. A. M. (2020). Categorization of green spaces for a sustainable environment and smart city architecture by utilizing big data. *Electronics*, *9*(6), 1028. https://doi.org/10.3390/ELECTRONICS9061028

Liu, Z. G., Li, X. Y., & Zhu, X. H. (2022). Scenario modeling for government big data governance decision-making: Chinese experience with public safety services. *Information & Management*, *59*(3), 103622. https://doi.org/10.1016/J.IM.2022.103622

Löfgren, K., & Webster, C. W. R. (2020). The value of big data in government: The case of 'smart cities.' *Big Data & Society*, *7*(1). https://doi.org/10.1177/2053951720912775

Lv, Z., Qiao, L., Cai, K., & Wang, Q. (2021). Big data analysis technology for electric vehicle networks in smart cities. *IEEE Transactions on Intelligent Transportation Systems*, *22*(3), 1807–1816. https://doi.org/10.1109/TITS.2020.3008884

Mohamed, N., Al-Jaroodi, J., & Jawhar, I. (2018). Service-oriented big data analytics for improving buildings energy management in smart cities. *2018 14th International Wireless Communications and Mobile Computing Conference*, 1243–1248. https://doi.org/10.1109/IWCMC.2018.8450469

Mohammadi, M., Student Member, G., Al-Fuqaha, A., & Member, S. (2018). Enabling cognitive smart cities using big data and machine learning: Approaches and challenges. *IEEE Communications Magazine*, *56*(2), 94–101. https://ieeexplore.ieee.org/abstract/document/8291121/

Mora, L., Deakin, M., & Reid, A. (2019). Combining co-citation clustering and text-based analysis to reveal the main development paths of smart cities. *Technological Forecasting and Social Change*, *142*, 56–69. https://doi.org/10.1016/J.TECHFORE.2018.07.019

Moreno, M. V., Terroso-Saenz, F., Gonzalez-Vidal, A., Valdes-Vela, M., Skarmeta, A. F., Zamora, M. A., & Chang, V. (2017). Applicability of big data techniques to smart cities deployments. *IEEE Transactions on Industrial Informatics*, *13*(2), 800–809. https://doi.org/10.1109/TII.2016.2605581

Mujeeb, S., Javaid, N., Ilahi, M., Wadud, Z., Ishmanov, F., & Afzal, M. K. (2019). Deep long short-term memory: A new price and load forecasting scheme for big data in smart cities. *Sustainability*, *11*(4), 987. https://doi.org/10.3390/SU11040987

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

Mystakidis, A., & Tjortjis, C. (2020). Big data mining for smart cities: Predicting traffic congestion using classification. *11th International Conference on Information, Intelligence, Systems and Applications, IISA 2020*, 1–8. https://doi.org/10.1109/IISA50023.2020.9284399

Nandury, S. V., & Begum, B. A. (2018). Big data for smart grid operation in smart cities. *2017 International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2017*, 1507–1511. https://doi.org/10.1109/WISPNET.2017.8300013

Naoui, M. A., Lejdel, B., Ayad, M., Amamra, A., & kazar, O. (2021). Using a distributed deep learning algorithm for analyzing big data in smart cities. *Smart and Sustainable Built Environment*, *10*(1), 90–105. https://doi.org/10.1108/SASBE-04-2019-0040/FULL/XML

Navarro, J. M., Tomas-Gabarron, J. B., & Escolano, J. (2017). A big data framework for urban noise analysis and management in smart cities. *Acta Acustica United with Acustica*, *103*(4), 552–560. https://doi.org/10.3813/AAA.919084

Nica, E., Konecny, V., Poliak, M., & Kliestik, T. (2020). Big data management of smart sustainable cities: Networked digital technologies and automated algorithmic decision-making processes. *Management Research and Practice*, *12*(2), 48–57.

Obinikpo, A. A., & Kantarci, B. (2017). Big sensed data meets deep learning for smarter health care in smart cities. *Journal of Sensor and Actuator Networks*, *6*(4), 26. https://doi.org/10.3390/JSAN6040026

Osman, A. M. S. (2019). A novel big data analytics framework for smart cities. *Future Generation Computer Systems*, *91*, 620–633. https://doi.org/10.1016/J.FUTURE.2018.06.046

Pereira, J., Ricardo, L., Luís, M., Senna, C., & Sargento, S. (2019). Assessing the reliability of fog computing for smart mobility applications in VANETs. *Future Generation Computer Systems*, *94,* 317–332. https://doi.org/10.1016/J.FUTURE.2018.11.043

Pérez-Chacón, R., Luna-Romera, J. M., Troncoso, A., Martínez-Alvarez, F., & Riquelme, J. C. (2018). Big data analytics for discovering electricity consumption patterns in smart cities. *Energies*, *11*(3), 683. https://doi.org/10.3390/EN11030683

Pramanik, M. I., Lau, R. Y. K., Demirkan, H., & Azad, M. A. K. (2017). Smart health: Big data enabled health paradigm within smart cities. *Expert Systems with Applications*, *87*, 370–383. https://doi.org/10.1016/J.ESWA.2017.06.027

Qiu, Y., Li, X., Zheng, W., Hu, Q., Wei, Z., & Yue, Y. (2017). The prediction of the impact of climatic factors on short-term electric power load based on the big data of smart city. *Journal of Physics: Conference Series*, *887*(1), 012023. https://doi.org/10.1088/1742-6596/887/1/012023

Rani, S., & Chauhdary, S. H. (2018). A novel framework and enhanced QoS big data protocol for smart city applications. *Sensors*, *18*(11), 3980. https://doi.org/10.3390/S18113980

Rathore, M. M., Paul, A., Ahmad, A., & Jeon, G. (2017). IoT-based big Data: From smart city towards next generation super city planning. *International Journal on Semantic Web and Information Systems*, *13*(1), 28–47. https://doi.org/10.4018/IJSWIS.2017010103

Rathore, M. M., Paul, A., Hong, W. H., Seo, H. C., Awan, I., & Saeed, S. (2018). Exploiting IoT and big data analytics: Defining Smart Digital City using real-time urban data. *Sustainable Cities and Society*, *40,* 600–610. https://doi.org/10.1016/J.SCS.2017.12.022

Sayah, Z., Kazar, O., Lejdel, B., Laouid, A., & Ghenabzia, A. (2020). An intelligent system for energy management in smart cities based on big data and ontology. *Smart and Sustainable Built Environment*, *10*(2), 169–192. https://doi.org/10.1108/SASBE-07-2019-0087/FULL/XML

Schatzinger, S., & Lim, C. Y. R. (2017). Taxi of the future: Big data analysis as a framework for future urban fleets in smart cities. *Smart and Sustainable Planning for Cities and Regions*, *2015*, 83–98. https://doi.org/10.1007/978-3-319-44899-2_6/COVER/

Shams, S., Goswami, S., Lee, K., Yang, S., & Park, S. J. (2018). Towards distributed cyberinfrastructure for smart cities using big data and deep learning technologies. *International Conference on Distributed Computing Systems*, 1276–1283. https://doi.org/10.1109/ICDCS.2018.00127

Sharif, A., Li, J., Khalil, M., Kumar, R., Sharif, M. I., & Sharif, A. (2017). Internet of things - Smart traffic management system for smart cities using big data analytics. *13th International Computer Conference on Wavelet Active Media Technology and Information Processing, ICCWAMTIP 2017*, 281–284. https://doi.org/10.1109/ICCWAMTIP.2017.8301496

Sharma, R., Gupta, D., Maseleno, A., & Peng, S.-L. (2022). Introduction to the special issue on big data analytics with internet of things-oriented infrastructures for future smart cities. *Expert Systems*, *39*(5), e12969. https://doi.org/10.1111/EXSY.12969

Shin, H. (2020). Analysis of subway passenger flow for a smarter city: Knowledge extraction from Seoul metro's "untraceable" big data. *IEEE Access*, *8*, 69296–69310. https://doi.org/10.1109/ACCESS.2020.2985734

Silva, B. N., Khan, M., & Han, K. (2020). Integration of big data analytics embedded smart city architecture with RESTful web of things for efficient service provision and energy management. *Future Generation Computer Systems*, *107*, 975–987. https://doi.org/10.1016/J.FUTURE.2017.06.024

Soomro, K., Bhutta, M. N. M., Khan, Z., & Tahir, M. A. (2019). Smart city big data analytics: An advanced review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(5), e1319. https://doi.org/10.1002/WIDM.1319

Tian, Z., Li, X., Wang, S., -, al, Zhang, B., Sun, Y.-H., Han, W., Chen, R., Xiong, Y., Li, J., Wei, Z., Li, X., Li, X., Hu, Q., Zhang, H., & Cui, P. (2017). Medium- and long-term electric power demand forecasting based on the big data of smart city. *Journal of Physics: Conference Series*, *887*(1), 012025. https://doi.org/10.1088/1742-6596/887/1/012025

Tiwari, P., Ilavarasan, P. V., & Punia, S. (2019). Content analysis of literature on big data in smart cities. *Benchmarking*, *28*(5), 1837–1857. https://doi.org/10.1108/BIJ-12-2018-0442/FULL/XML

van Eck, N. J., & Waltman, L. (2017). Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics*, *111*(2), 1053–1070. https://doi.org/10.1007/S11192-017-2300-7/TABLES/4

Wang, T., Bhuiyan, M. Z. A., Wang, G., Rahman, M. A., Wu, J., & Cao, J. (2018). Big data reduction for a smart city's critical infrastructural health monitoring. *IEEE Communications Magazine*, *56*(3), 128–133. https://doi.org/10.1109/MCOM.2018.1700303

Wang, W. (2020). The research on street landscape design in smart city based on big data. *Advances in Intelligent Systems and Computing*, *928*, 1327–1331. https://doi.org/10.1007/978-3-030-15235-2_181/COVER/

Wei, L., & Yang, S. (2018). Based on big data technology analysis on the mode and countermeasures of smart city construction operation management. *3rd International Conference on Smart City and Systems Engineering, ICSCSE 2018*, 184–186. https://doi.org/10.1109/ICSCSE.2018.00043

Xiao, X., & Xie, C. (2021). Rational planning and urban governance based on smart cities and big data. *Environmental Technology & Innovation*, *21*, 101381. https://doi.org/10.1016/J.ETI.2021.101381

Xie, H., Shi, W., Choudhary, H., Fu, H., & Guo, X. (2019). Big data analysis for retrofit projects in smart cities. *2019 3rd International Conference on Smart Grid and Smart Cities, ICSGSC 2019*, 1–5. https://doi.org/10.1109/ICSGSC.2019.00-28

Yang, C., Su, G., & Chen, J. (2017). Using big data to enhance crisis response and disaster resilience for a smart city. *2017 IEEE 2nd International Conference on Big Data Analysis, ICBDA 2017*, 504–507. https://doi.org/10.1109/ICBDA.2017.8078684

Yang, Z., Jianjun, L., Faqiri, H., Shafik, W., Talal Abdulrahman, A., Yusuf, M., & Sharawy,

A. M. (2021). Green internet of things and big data application in smart cities development. *Complexity*, *2021*, 4922697. https://doi.org/10.1155/2021/4922697

Zaree, T., & Honarvar, A. R. (2018). Improvement of air pollution prediction in a smart city and its correlation with weather conditions using metrological big data. *Turkish Journal of Electrical Engineering and Computer Sciences*, *26*(3), 1302–1313. https://doi.org/10.3906/elk-1707-99

Zhang, Y., Geng, P., Sivaparthipan, C. B., & Muthu, B. A. (2021). Big data and artificial intelligence based early risk warning system of fire hazard for smart cities. *Sustainable Energy Technologies and Assessments*, *45*, 100986. https://doi.org/10.1016/J.SETA.2020.100986

Zhao, Z., & Zhang, Y. (2020). Impact of smart city planning and construction on economic and social benefits based on big data analysis. *Complexity*, *2020*, 8879132. https://doi.org/10.1155/2020/8879132

Zhou, J., Wang, P., & Xie, L. (2020). Research on resource allocation optimization of smart city based on big data. *IEEE Access, 8*, 158852–158861. https://doi.org/10.1109/ACCESS.2020.3017765

Zhu, L., Yu, F. R., Wang, Y., Ning, B., & Tang, T. (2019). Big data analytics in intelligent transportation systems: A survey. *IEEE Transactions on Intelligent Transportation Systems*, *20*(1), 383–398. https://doi.org/10.1109/TITS.2018.2815678