



## Big Data in Smart City

Matthew N. O. Sadiku<sup>1</sup>, Uwakwe C. Chukwu<sup>2</sup>, Abayomi Ajayi-Majebi<sup>3</sup>,  
Sarhan M. Musa<sup>1</sup>

<sup>1</sup>Roy G. Perry College of Engineering, Prairie View A&M University, Prairie View, TX, USA

<sup>2</sup>Department of Engineering Technology, South Carolina State University, Orangeburg, SC, USA

<sup>3</sup>Department of Manufacturing Engineering, Central State University, P.O. Box 1004, Wilberforce, OH, USA

Email: sadiku@ieee.org; uchukwu@scsu.edu; ajayi-majebi@centralstate.edu;

smmusa @pvamu.edu

**Abstract** Every large urban area is being identified as a future smart city. With 68 per cent of the global population set to live in urban areas by 2050, there is a push to make the world's cities more connected, efficient, and citizen-friendly. Smart cities will change our lives at many levels such as less pollution, better parking, efficient lighting, smart energy grid, improved traffic flow, smarter security, improved waste management, improved transportation networks, and greener environment. The proper use of big data can revolutionize the smart city initiatives. This paper explains how big data is used in helping cities provide new services and improve quality of life for the residents.

**Keywords** big data, big data analytics, smart city

### Introduction

More people are choosing to live in cities more than ever before. Major cities are experiencing some urbanization challenges which include managing power usage efficiently, transportation problems, environmental pollution, resources shortage, and traffic jam. Smart city initiative aims to mitigate the challenges by the integration of municipal service, business, transportation, water, energy source and other urban sub-systems through close combination of human wisdom and ICT infrastructure (such as Internet of things connected sensors) [1].



Figure 1: Components of a smart city [2].



Its goal is to improve the quality of life of the citizens by providing intelligent services in a wide variety of aspects like transportation, healthcare, education, environment, water, and energy. Figure 1 shows the various components of a smart city [2].

Cities all over the world are becoming smarter, greener, and safer urban environments, with cleaner air and water, better mobility, and efficient public services. For a city to become smart, it must use technology and data with the goal of delivering a better quality of life of its residents. By definition, smart cities are data-driven. They depend on reliable data to base their long-term decisions on. Data sharing (public and private sharing) is a major requirement for smart city deployments. It will improve privacy, interoperability, and security. Smart cities all utilize sensor data to improve city management. This is where big data and analytics play an important role. Big data can help cities to monitor and manage issues such as waste disposal, transportation, and saving resources. This makes big data an integral part of creating and maintaining smart cities.

### Big Data Characteristics

Big data (BD) is a relatively newer technology that can make use of smart city services. The three main sources of big data are machines, people, and companies. As shown in Figure 2 [3], big data is often described with the five “Vs”: volume, velocity, variety, veracity, and value [4].

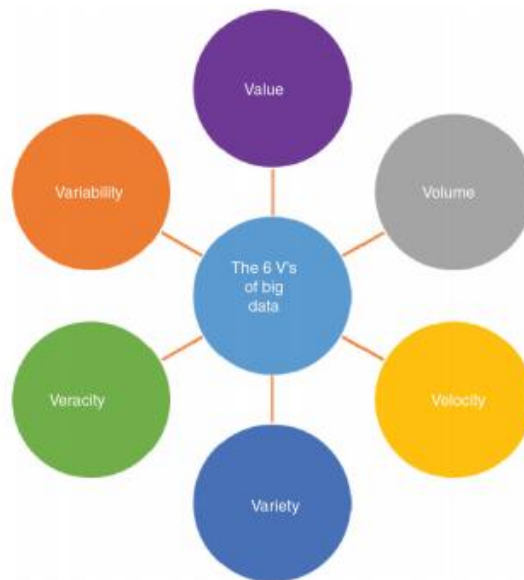


Figure 2: Big data's 6V [3]

- **Volume:** This refers to the size of the data being generated both inside and outside organizations and is increasing annually. Some regard big data as data over one petabyte in volume.
- **Velocity:** This depicts the unprecedented speed at which data are generated by Internet users, mobile users, social media, etc. Data are generated and processed in a fast way to extract useful, relevant information. Big data could be analyzed in real time, and it has movement and velocity.
- **Variety:** This refers to the data types since big data may originate from heterogeneous sources and is in different formats (e.g., videos, images, audio, text, logs). BD comprises of structured, semi-structured or unstructured data.
- **Veracity:** By this, we mean the truthfulness of data, i.e. whether the data comes from a reputable, trustworthy, authentic, and accountable source. It suggests the inconsistency in the quality of different sources of big data. The data may not be 100% correct.
- **Value:** This is the most important aspect of the big data. It is the desired outcome of big data processing. It refers to the process of discovering hidden values from large datasets. It denotes the value derived from the analysis of the existing data. If one cannot extract some business value from the data, there is no use managing and storing it.



On this basis, small data can be regarded as having low volume, low velocity, low variety, low veracity, and low value. Additional five V's has been added [5]:

- *Validity*: This refers to the accuracy and correctness of data. It also indicates how up to date it is.
- *Viability*: This identifies the relevancy of data for each use case. Relevancy of data is required to maintain the desired and accurate outcome through analytical and predictive measures.
- *Volatility*: Since data are generated and change at a rapid rate, volatility determines how quickly data change.
- *Vulnerability*: The vulnerability of data is essential because privacy and security are of utmost importance for personal data.
- *Visualization*: Data needs to be presented unambiguously and attractively to the user. Proper visualization of large and complex clinical reports helps in finding valuable insights.

Instead of the 5V's above, some suggest the following 5V's: Venue, Variability, Vocabulary, Vagueness, and Validity) [6].

### Big Data Analytics

Every day, data is growing bigger and bigger, and big data analysis (BDA) has become a requirement for gaining invaluable insights into data such that companies could gain significant profits in the global market. Big data analytics and the Internet of Things (IoT) form the foundation of the smart city model. Smart city initiatives need big data analytics to function. Once the big data is ready for analysis, we use advanced software programs such as Hadoop, MapReduce, MongoDB, and NoSQL databases [7]. Big data analytics refers to how we can extract, validate, translate, and utilize big data as a new currency of information transactions. It is an emerging field that is aimed at creating empirical predictions. Data-driven organizations use analytics to guide decisions at all levels [8].

### Why Big Data in Smart Cities

The Internet of things (IoT) and big data can help significantly to convert a normal city into a smart city. Cities that are considered smart include London, Nice in France, Songdo in South Korea, and Santa Cruz. Smart cities are also known as intelligent cities, information cities, and virtual cities. A simple formula for smart city is as follows [9]:

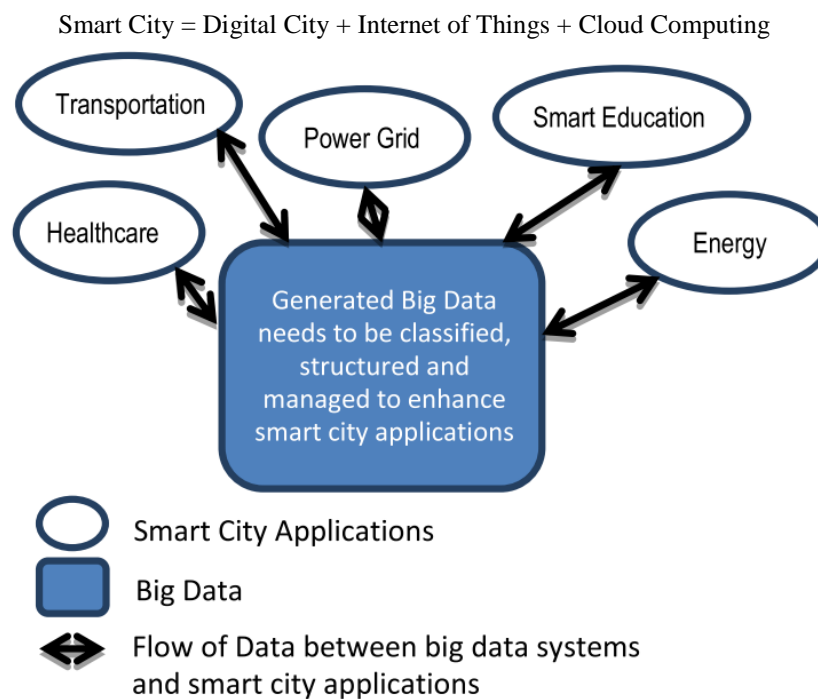


Figure 3: Relationship between big data and smart city [10]



A smart city typically has the following components -- mobility, governance, environment, and people-- as well as its applications and services such as healthcare, transportation, smart education, and smart grid. Some of the characteristics of smart city include sustainability, resilience, governance, enhanced quality of life of its citizens, and intelligent management of natural resources and city facilities. Data is being generated from multiple sources (such as smart phones, computers, sensors, cameras, GPS, people, etc.) within the city resulting in massive big data. Big data offers the potential for cities to obtain valuable insights from a large amount of data collected through various sources. It is essential to understanding how people in cities move, how energy is used, how various aspects of infrastructure interact, and much more. Figure 3 illustrates the relationship between big data and smart city [10].

In a smart city, big data plays an important role in processing data collected through IoT devices so that further analysis can be made to recognize the patterns and needs in the city.

### **Applications in Smart City**

The idea of smart city is widely perceived as a means of solving urban challenges which are regarded as wicked problems—intractable problems) through the integration of information technology (IT) with the city's infrastructure. Large corporations like IBM and Cisco are working with civic planning bodies, governments, and academic communities to develop data-based systems for waste management, transport, law enforcement, and energy. They are currently harnessing large quantities of big data to understand the complexities within cities and plan for the future.

Big data lies at the heart of smart city innovation. There are many areas where data-driven innovation is crucial. These include traffic management, public safety, infrastructure maintenance, and crowd control, smart education, smart traffic management, smart grid, smart healthcare, smart irrigation system [11,12]. We will briefly look at some of them.

- **Intelligent Traffic Management:** Cities are already harnessing data to deliver big improvements in urban mobility. For example, data collected from connected cars, cameras, IoT sensors, mobility apps, etc. can be used to make traffic flow more efficient, reduce congestion, and enable better traffic management. From cameras and sensors that can monitor parking availability, to traffic lights fed data in real time to facilitate more efficient traffic flow. Sensors in the parking lot would determine the best place to park.
- **Public Safety:** Smart city planners cannot afford to lose sight of the promise of big data for public safety. Public safety technology uses of GPS data to find missing persons. It deploys drones allowing for more effective search-and-rescue operations during major events. Smart city public safety solutions can reduce fatalities, crime incidences, and emergency response times. It is easier for civilians to report crimes to security agencies. Big data could allow for a comprehensive study of crime in focused areas as well as implement a preventive model to predict future patterns of crime.
- **Managing Infrastructure:** Population growth and climate change affect cities' infrastructure. Drones can monitor hard-to-reach areas and generate actionable data to guide decisions on infrastructure repair and maintenance. Supply chains depend on good infrastructure: roads, bridges, trains, electric grid, etc.
- **Crowd Control:** As the world's population increases and cities become more crowded, crowd control becomes necessary. Whenever a large group of people congregates in one place, there may be an unplanned need for meeting the crowd's needs such as food, drink, security, emergency responders, etc. Large congregation of people translates to massive data. This is an area that big data is already being used to great success. Big data is being employed to understand when, how, and why crowds form, and to predict their movements and actions. **For example**, every year, Saudi Arabia's population increases by more than three million people visiting Mecca for the five days of the Hajj. To avoid tragedies Saudi Arabia is incorporating numerous ideas suggested by big data to control such crowds [13].
- **Water Management:** Real-time analysis and sensors can help detect the flow of water, pollution level, monitor usage, reduce areas of leakage, sewage overflow, etc.



- **Transport Management:** This is a major problem faced by cities across the world. Big data can play an important role in transport management. Information on public buses and other modes of transport is available in your smart phone.
- **Smart Mobility and Transportation:** With increasing overcrowding in cities, transportation will play a key role in decongesting smart cities of the future.
- **Smart Governance:** Smart cities are supposed to enhance the governance of urban areas via the employment of data. Big data can help develop stronger bridges between citizens and the government. It makes public policymaking more citizen-focused. It enables a wise management of natural resources, through participatory governance. It also makes it easy to use data to recognize the patterns and needs in the city. Governments can address local issues faster and create a citizen-centered model of governance that increases transparency.

Other areas where big data is being used in smart city include public health, crime prevention, efficient spending, pollution control, garbage management, infrastructure planning, health monitoring, education planning, and smart burglar alarm.

### **Benefits and Challenges**

Governments and researchers are currently harnessing large quantities of big data to understand the complexities within cities and anticipate future. Big data can contribute to smart cities in the following ways [14]:

- The traffic will be measured and regulated with the help of RFID tags on the vehicles.
- Even garbage collection will generate data. Each house will have garbage disposal units and garbage will be sucked from them, eliminating the need for garbage trucks.
- Data will make life more secured for the citizens.
- The smart energy grid can control the street lights.
- Big data can help reduce carbon emissions and bring down pollution.
- Parking problems can be better managed. Cars will have sensors attached which can guide the car to the nearest available parking lots.
- The environment will be cooler and greener with less energy being consumed.
- Ensuring that cities retain high levels of liveability is imperative.
- Can enable truly evidence-based, data-driven decision making.
- Many new opportunities exist for social interaction and more informed decision-making.

Without doubts big data is potentially able to make enormous contributions to the development of smart cities. However, the smart city vision faces some challenges which should not go unaddressed. The challenges include the following [14]:

- Some critics think that combining smart city and big data will not work.
- Ensuring the safety of citizens is one of top priorities for every city.
- Security and data confidentiality are major issues.
- Privacy is an issue since smart cities deal with highly personal information.
- Availability of funds, data confidentiality and social issues are important.
- The issue of affordability in third world nations needs to be addressed.
- Huge data from city surveillance cameras poses challenges of transmission, storage, and analysis.
- While enormous data is collected, traditional barriers hinder the needed sharing, analysis, and interpretation of the data among different city departments.

### **Conclusion**

Smart city and big data are relatively and important concepts. We live in the big data era. Big data consists of huge amounts of structured and unstructured data. BDA tools provide a means of analyzing such data and drawing conclusions about them to help organizations and governments make informed decisions. As a result, many governments around the world have started to utilize big data to support the development and sustainability of smart cities.



Cities around the world are using technology to transform the way the residents live, becoming connected smart cities, and improving the living standards. Smart cities are part of a future trend in urban design and are key to solving urban problems. More information about the use of big data in smart city can be found in books in [15-22] and the following related journals:

- *Journal of Big Data*
- *Big Data Research*
- *Journal of Internet Services and Applications*

## References

- [1]. Y. Wu et al., "Smart city with Chinese characteristics against the background of big data: Idea, action and risk," *Journal of Cleaner Production*, vol. 173, February 2018, pp. 60-66.
- [2]. "How IOT and big data are driving smart traffic management and smart cities," <https://sileotech.com/how-iot-and-big-data-are-driving-smart-traffic-management-and-smart-cities/>
- [3]. B. Ristevski and M. Chen, "Big data analytics in medicine and healthcare," *Journal of Integrative Bioinformatics*, 2018.
- [4]. M. N.O. Sadiku, M. Tembely, and S.M. Musa, "Big data: An introduction for engineers," *Journal of Scientific and Engineering Research*, vol. 3, no. 2, 2016, pp. 106-108.
- [5]. P/ K. D. Pramanik S. Pal, and M. Mukhopadhyay, "Healthcare big data: A comprehensive overview," in N. Bouchemal (ed.), *Intelligent Systems for Healthcare Management and Delivery*. IGI Global, chapter 4, 2019, pp. 72-100.
- [6]. J. Moorthy et al., "Big data: Prospects and challenges," *The Journal for Decision Makers*, vol. 40, no. 1, 2015, pp. 74–96. <https://www.grandviewresearch.com/industry-analysis/industrial-wireless-sensor-networks-iwsn-market>
- [7]. M. N. O. Sadiku, J. Foreman, and S. M. Musa, "Big data analytics: A primer," *International Journal of Technologies and Management Research*, vol. 5, no. 9, September 2018, pp. 44-49.
- [8]. C. M. M. Kotteti, M. N. O. Sadiku, S. M. Musa, "Big data analytics," *Invention Journal of Research Technology in Engineering & Management*, vol. 2, no. 10, Oct. 2018, pp. 2455-3689.
- [9]. L. DeRen, J. J. Cao, Y. Yao, "Big data in smart cities," *Science China Information Sciences*, vol 58, October 2015.
- [10]. E. A. Nuaimi et al., "Applications of big data to smart cities," *Journal of Internet Services and Applications*, vol. 6, no. 25, 2015.
- [11]. A. Kaplan, "How big data will drive smart city innovation," August 2019, <https://www.itproportal.com/features/how-big-data-will-drive-smart-city-innovation/>
- [12]. "Understanding how analytics can be useful for developing smart city," <https://www.proschoolonline.com/blog/how-analytics-can-be-useful-for-developing-smart-cities>
- [13]. D. Newman, "Big data and the future of smart cities," August 2016, <https://www.forbes.com/sites/danielnewman/2016/08/15/big-data-and-the-future-of-smart-cities/#5968f19b26b8>
- [14]. K. Pal, "How big data helps build smart cities," <https://www.kdnuggets.com/2015/10/big-data-smart-cities.html>
- [15]. H. Liu, *Smart Cities: Big Data Prediction Methods and Applications*. Singapore: Springer, 2020.
- [16]. B. McQuee, *Big Data Analytics for Connected Vehicles and Smart Cities*. Artech House, 2017.
- [17]. M. Sarsar, *The Big Disruption: IoT, Smart Cities, Connected Vehicles, Big Data Analytics: A Practitioner Point of View and Return of Experience (REX)*. Amazon, 2019.
- [18]. A. M. Townsend, *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia*. W. W. Norton & Company, 2013.
- [19]. K. Danieli, *Monetization of Big Data in Smart Cities: An Analytical View on Dta Monetization in Switzerland*. Universität St. Gallen, 2019.
- [20]. A. Alavi and W. G. Buttler, *Data Analytics for Smart Cities*. Boca Raton, FL: CRC Press, 2018.



- [21]. M. Atiquzzaman, N. Yen, and Z. Xu (eds.), Big Data Analytics for Cyber-Physical System in Smart City. Springer, 2020.
- [22]. X. Liu et al., Big Data and Smart Service Systems. Academic Press, 2016.

#### About Authors

**Matthew N.O. Sadiku** is a professor emeritus in the Department of Electrical and Computer Engineering at Prairie View A&M University, Prairie View, Texas. He is the author of several books and papers. His areas of research interest include computational electromagnetics and computer networks. He is a fellow of IEEE.

**Uwakwe C. Chukwu** is an associate professor in the Department of Industrial & Electrical Engineering Technology of South Carolina State University. He has published several books and papers. His research interests are power systems, smart grid, V2G, energy scavenging, renewable energies, and microgrids.

**Abayomi Ajayi-Majebi** is a professor in the Department of Manufacturing Engineering at Central State University in Wilberforce, Ohio. In 2015 he was honored by the White House as a Champion of Change for his significant contributions to the engineering education of minority students. He is a senior member of both the Society of Manufacturing Engineers and the American Society for Quality.

**Sarhan M. Musa** is a professor in the Department Electrical and Computer Engineering at Prairie View A&M University, Texas. He has been the director of Prairie View Networking Academy, Texas, since 2004. He is an LTD Sprint and Boeing Welliver Fellow. His areas of research interest include computational electromagnetics and computer networks.

