

A Survey on Big Data Analytics Architecture for Urban Transportation System: A Case for Nairobi Metropolitan

Justin Muchiri Njeru
School of Computing and Informatics
University of Nairobi, Kenya

Elisha Odira Abade
School of Computing and Informatics
University of Nairobi, Kenya

ABSTRACT

In the recent past, studies in the Intelligent Transportation Systems are on high gear to bring a solution to urban traffic systems caused by high rate of urbanization in most cities. Urbanization in Nairobi City has witnessed rapid growth over the last 10 years. Researchers have found that expansion of road networks does not solve the traffic situations. There is great loss of productivity for commuters who have to spend long hours in traffic. This research found that traffic information on updates, congestion and incidents is shared in various formats in various channels including radio, google maps and social media platforms. This information is big data and therefore need to establish architecture to analysis and provide reports for road users.

Keywords

Architecture of Big Data Analytics, Urban Transportation Systems, Big Data Sources

1. INTRODUCTION

Urban Transportation Systems are very important in city life; however, in developing countries the transportation systems are characterized by congestion, chaos and gridlocks. In the recent past, studies in the Intelligent Transportation Systems (Bekiaris, 2017), (International Transport Forum, 2019) have found that expansion of road networks does not solve the traffic situations. Analysis of big data has profound impact on the way data is used for City planning, traffic management, accident reporting dashboard, predicting flow of traffic, enhanced travel plans and distribution of social amenities (Zhu et al., 2019). Unfortunately, other international transportation models have failed to work in Nairobi scenarios due to cultural, historical planning challenges and economic aspects (Klopp et al., 2015). Analyses of heterogeneous data offers a sure step to understand, automate and provide insights for decision making in urban transportation systems (Robert et al., 2016) and (Lee, 2019).

2. RELATED WORKS

Several techniques are used to control traffic in different cities; traffic lights systems, fuser logic and the analysis has shown their limitations (Daiheng, 2016). Effects of legislation of Laws, Guidelines, Standards and Principles for protection of data, is an area of concern to both road users and policy makers. This study has reviewed the applicability of the relevant principles in promotion of the research and development of the Big Data related projects (Korbel, 2019).

2.1 Big Data Architecture

Big data architecture is the overarching system used to receive and process enormous amounts of data from various sources and it's often referred to as "big data" (NIST, 2019). The

architecture can be considered the blueprint for a big data solution based on the types of processing activities. Big data architecture is designed to handle the following types of work: Transmission, Reception, Storage, Manipulation and Retrieval, such as Hadoop (Silva et al, 2018).

2.2 Artificial Intelligence in Transportation

A negative influence of Big Data impacts on road safety, for instance, drivers might be distracted from driving by new services, from navigation and infotainment to Cooperative Intelligent Transport Systems (C-ITS) alerts (Barrachina et al., 2015). It is important for transport players to determine where, how and when to use these technologies to make a rapid improvement in relieving congestion, making travel time informative security & safety bulletins through Twitter, Google and Facebook.

2.3 Digital Matatus and Ma3Route Projects

The aim of Digital Matatus Project was bring a data standard for paratransit and flexible transportation services. The project has shown how to leverage on cellphone technology to collect data freely from road users for essential road infrastructure (Williams S, 2016; Deloitte, 2019). The project was conceived out of collaboration between University of Nairobi MIT Civic Data Design Lab and Columbia University through support of World Bank and Rockefeller Foundation (Sarah Williams, n.d.). The project found that indeed General Transit Feed Specifications (GTFS) can be adapted in Kenya and institutionalizing processes of data updating and sharing (GTFS, 2019; Columbus, 2016) (ILO, 2017). The GTFS project outcomes were as important as the data they produced since the researchers were able to translate data into paper-based traffic map. The outcome has provided data driven needs of information on traffic by commuters and also mobile applications have been developed from this research including; Ma3route, Flashcast, digitalmatatu, Kenyan Routes and Matatuapp (SASRA, 2019). The transit map has also been incorporated in Google maps app available on mobile and web based url following the success of the digital Matatu project.

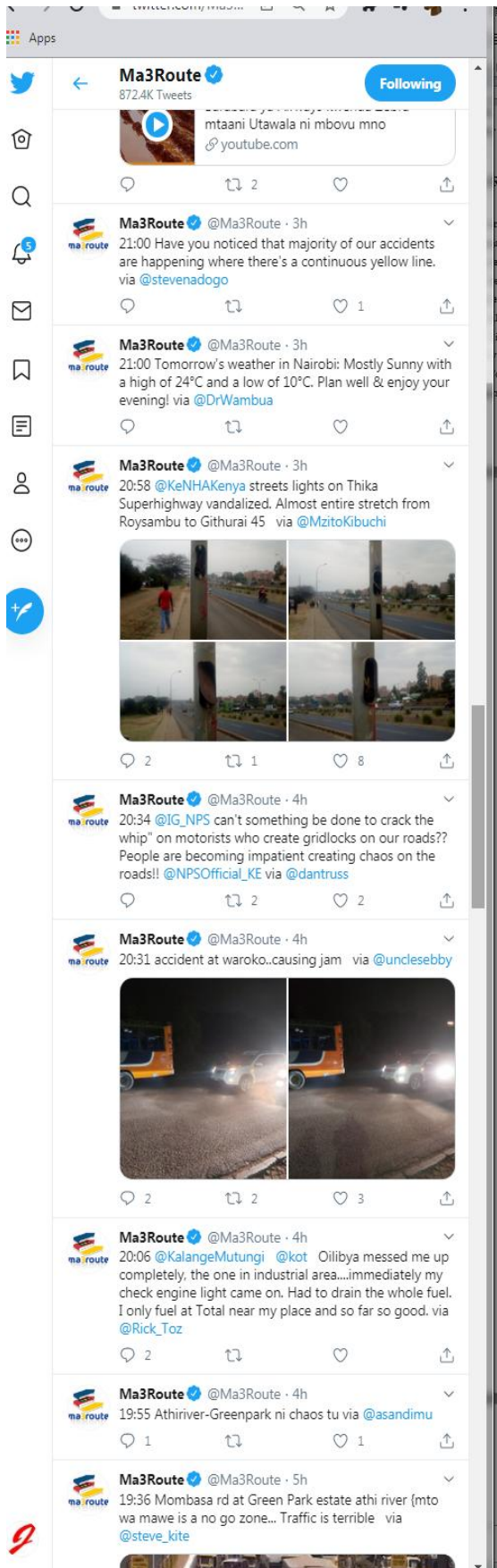


Figure 2: Twitter usage on Traffic Updates Credit: @Ma3Route

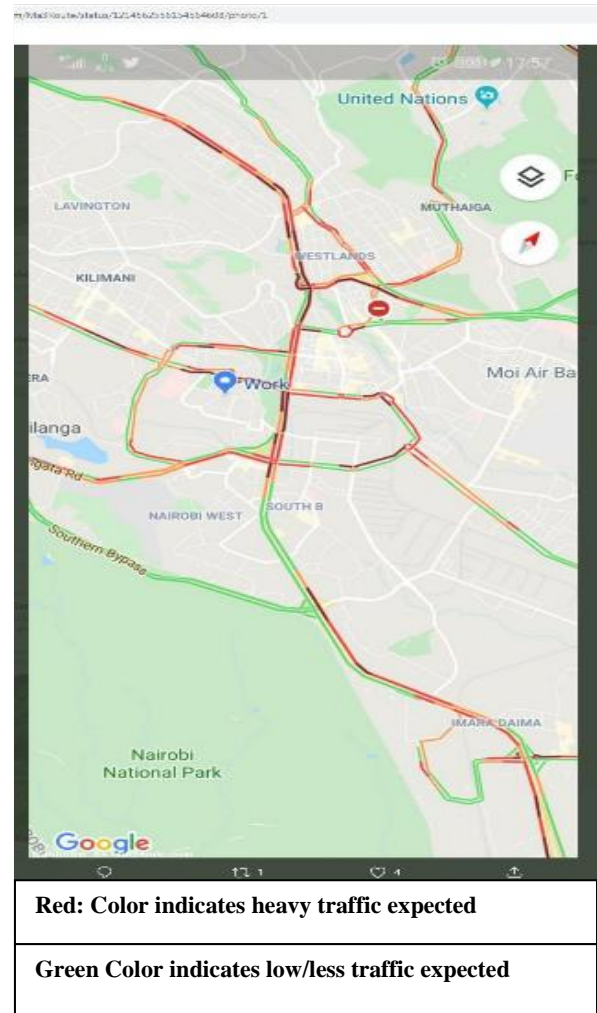


Figure 3: Google Maps indicating traffic Situation for a section of Nairobi City

3. CONCEPTUAL MODEL

To tests whether the objectives will be met by this research a conceptual model was adopted with independent variables comprising of; costs, technology, infrastructure, data, performance, skills and scalability. The controls used were adopted from the McKinsey's 7S model. The model visualizes why certain organizations or countries would implement a certain technology and succeed while others fail. These 7S moderating variables are: System, Skills, Staff, Style, Structure, Strategy and Shared Values. Despite that the McKinsey's model originated from the practice as opposed to theory, it has been proved as the ultimate model of testing the organization to achieve its objectives (Gokdeniz, 2017). It also represents the significance of interrelationships that exists among all the seven variables. However, the new information age requires for the model to be adjusted catering for emerging trends and needs (Oya, 2017).

Assessing social economic value of Big Data Analytics is an important aspect of knowing whether the venture is feasible to both investors, policy makers on researchers in this field (Hoti, 2015). This architecture has also identified important implementation stages that Nairobi City can utilize to achieve an intelligent transportation system through analyze of Big

Data. The architecture, however, also has to be extended to support behavior in a more complex manner than data collection (Hofman, 2015). Data governance rules and data policies will also require functionality like identification and

authentication, trust, privacy enhanced technology, etc. that is also required by complex systems. These types of architectural questions need further research.

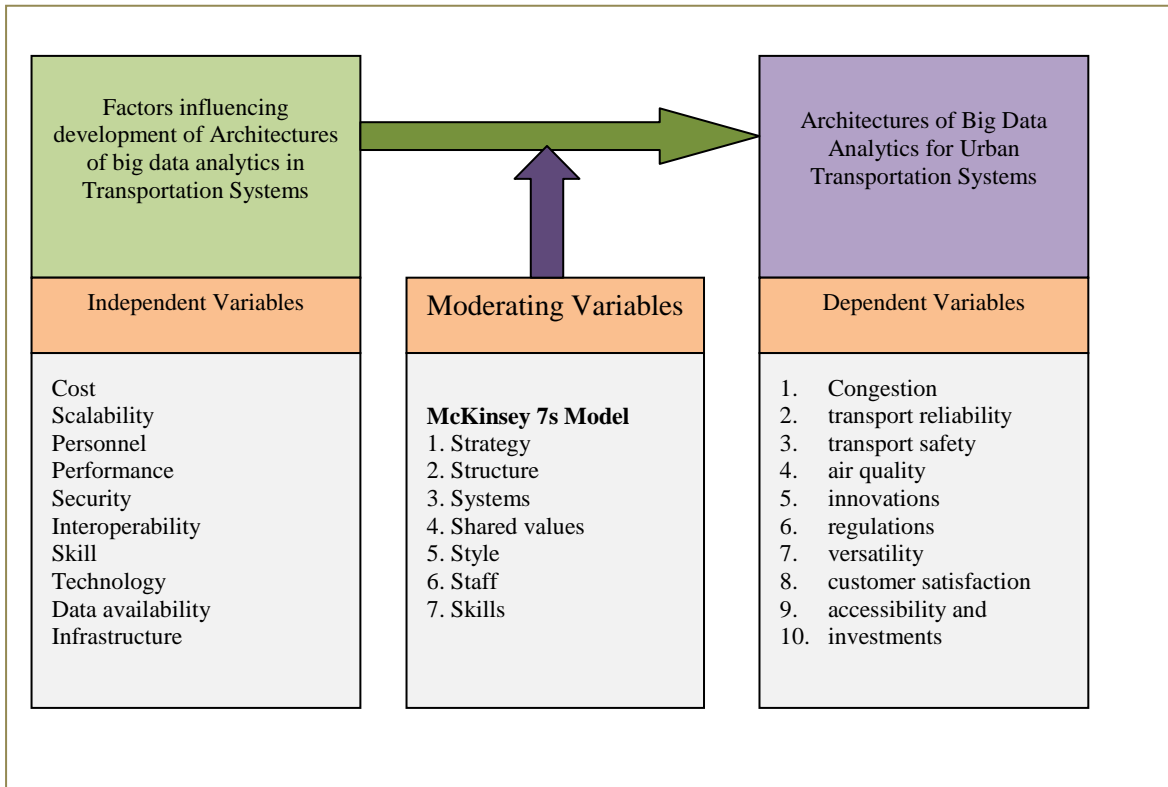


Figure 4: Conceptual Model

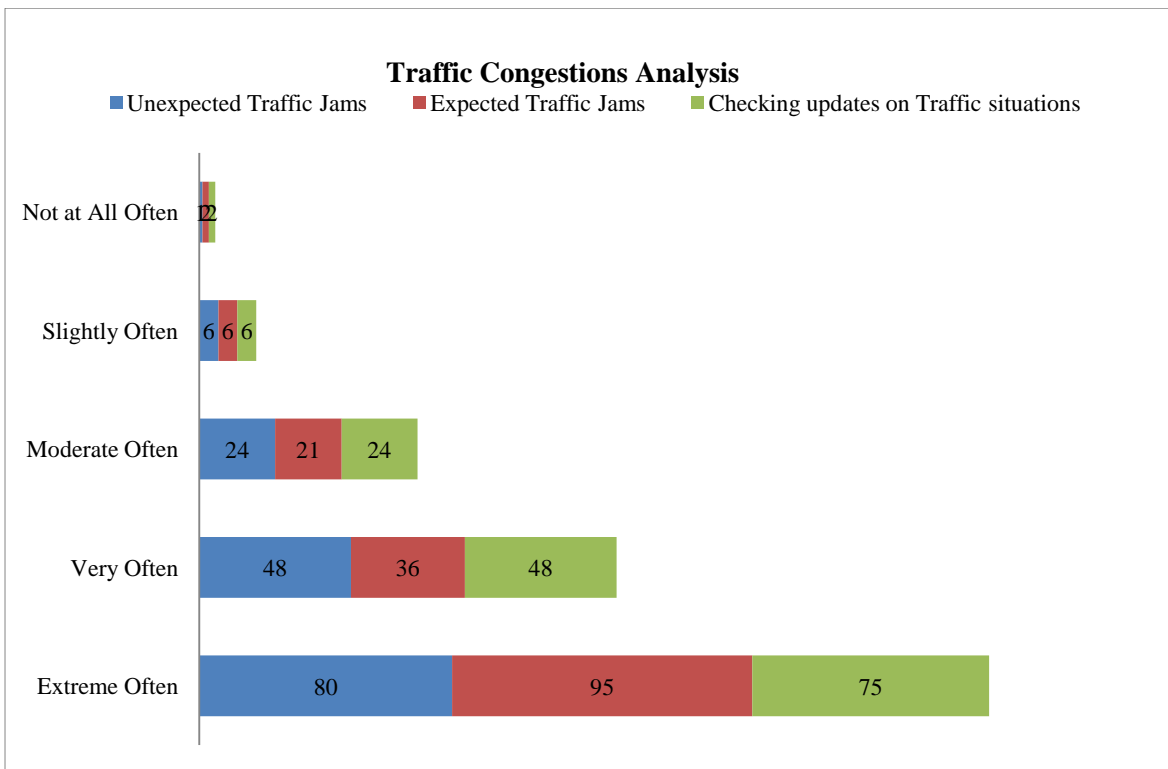


Figure 5: Traffic Congestion Analysis

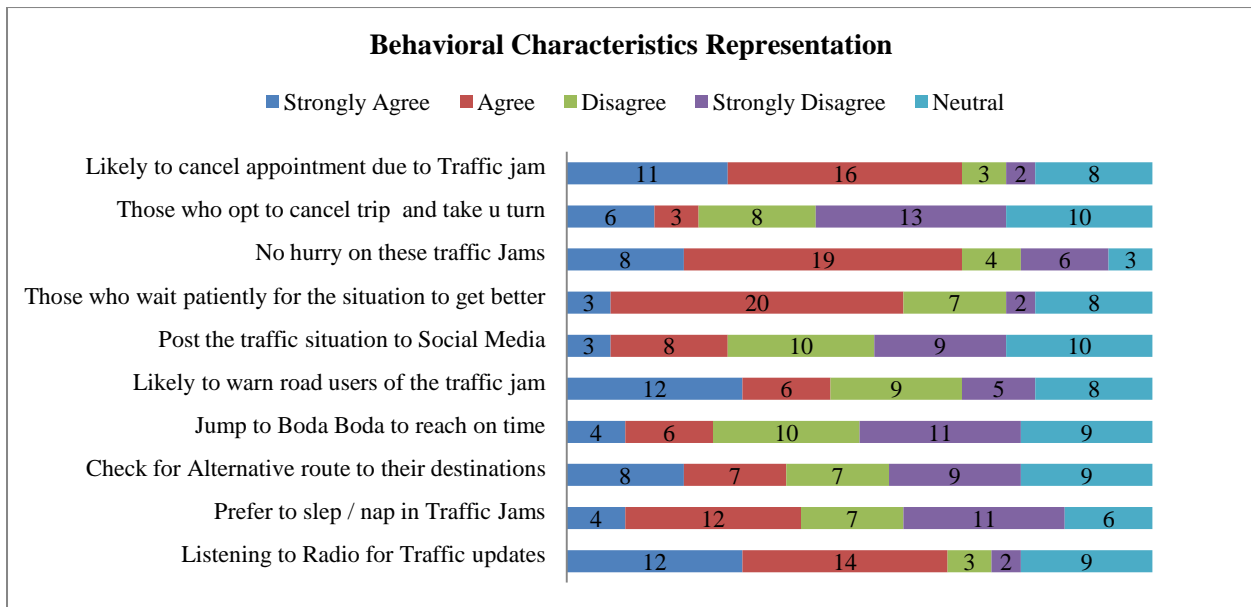


Figure 6: Behavioral representation of road users

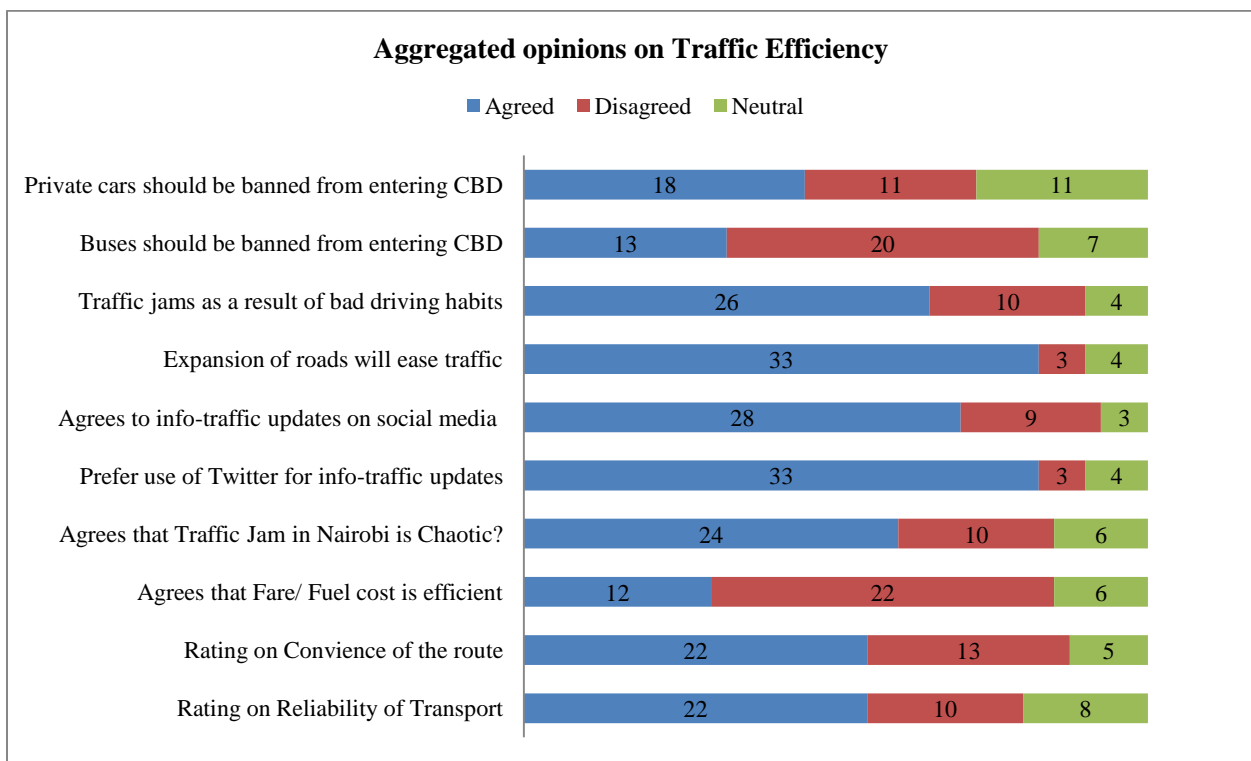


Figure 7: Aggregated opinions on Traffic Efficiency

The relationship of transport costs, insecurity, air pollution due to traffic was plotted against the reliability and customer satisfaction and insights indicated that they are indirectly proportional. The use of social media, access to 4/5G network and car radio traffic updates analysis as the emerging technologies associated to traffic updates relates to traffic congestion, and therefore the users prefer to get updates of the road situation as updated by other road users.

Regression analysis of how users can benefit from access to traffic updates from social media to enable them plan there

route accordingly indicates is measured to establish if there are gaps in information provided on social media. A 63% of the respondents indicated having access to social media and 4G network and 72% of this group trusted the information shared on the platforms. Route planning measures are emerging with 52% and are related to the choice of road user's preferences to access devices for updates on certain routes.

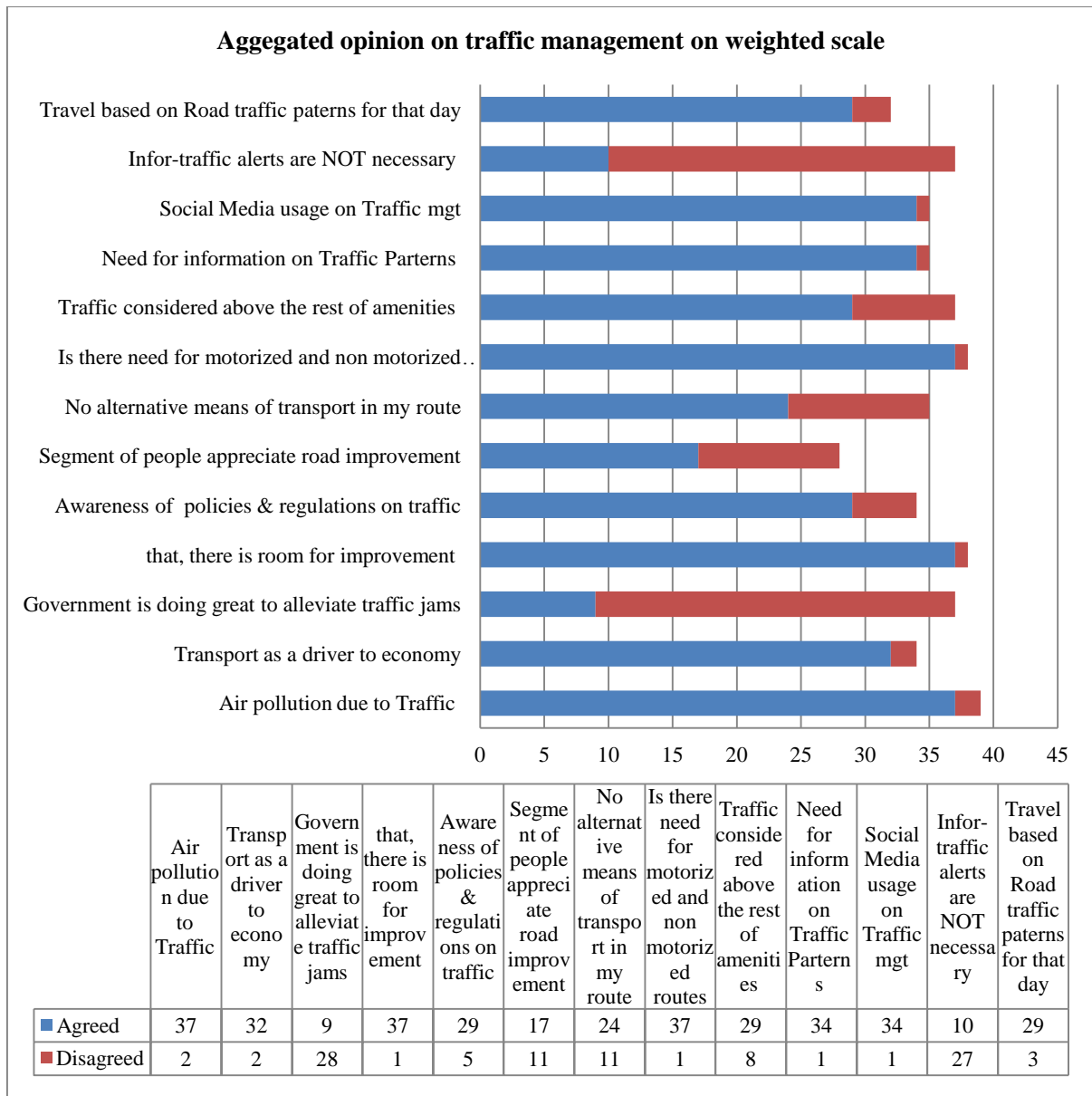


Figure 8: Aggregated opinions on road users on traffic management

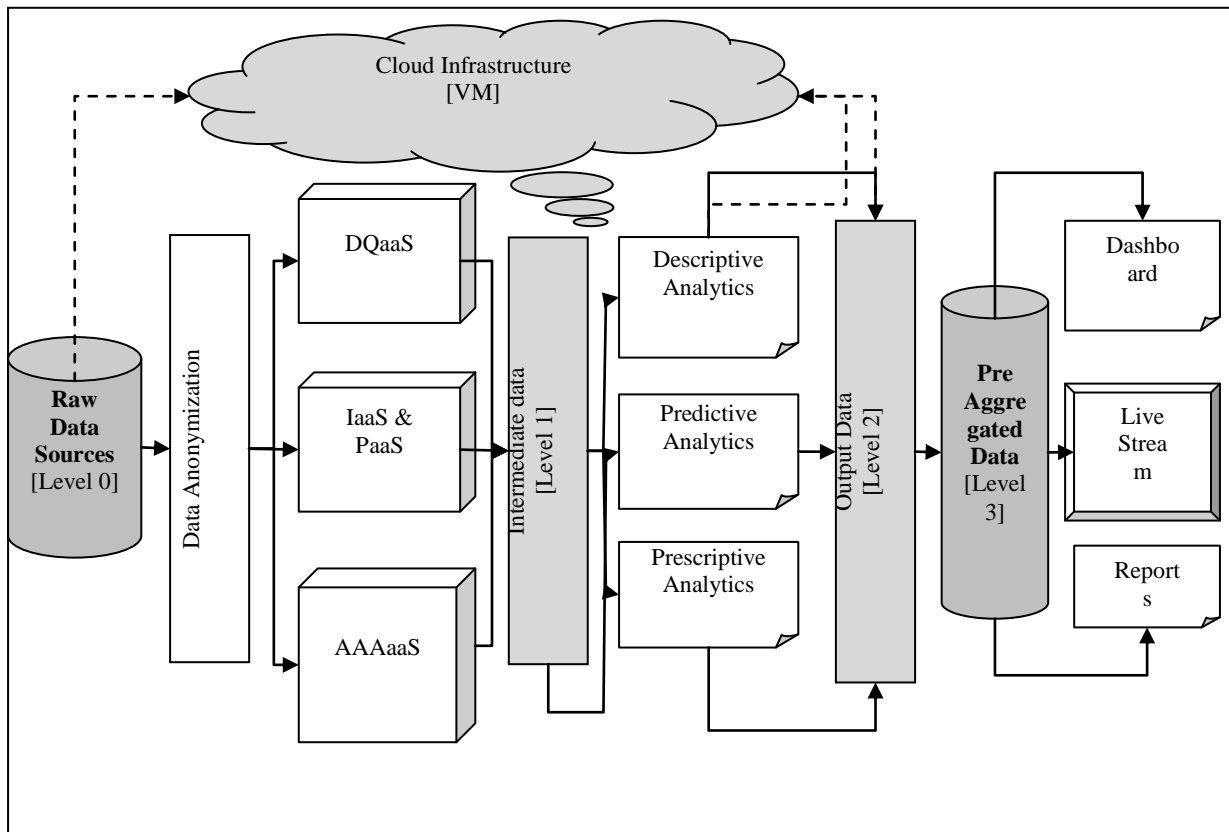


Figure 9: A Big Data Analytics Architecture

The aggregated opinions in the above table listed areas of concerns in traffic management and responses from the participants are as follows in summary:

- a) That traveler is cautiously planning for their journey based on traffic patterns of that day. Therefore, provision of reliable source of traffic information updates will change the situation by providing a central source of reliable traffic updates for the road users.
- b) Information on traffic has been listed as necessary by road users
- c) That, Social media as the central focal point for accessing traffic updates as well as socializing app has been appreciated and adopted by over 90% of the respondents (KNBS , 2018).
- d) Need for motorized and non-motorized means of transportation is a requirement for ensuring mobility thrives in Nairobi city. This is a consideration of foot paths, bicycle lanes and in physical road planning.
- e) Over 90% of the respondent's indicated that there is need to improve road infrastructure while indicating lower rating that government has done less to their expectation on road improvement programs (NAMATA, 2017) (UN Habitat,, 2019).
- f) Contribution of air pollution can be attributed to traffic congestion according to the results of the survey. This is a great concern for policy makers in matters of environment and health.

4. CONCLUSION AND RECOMMENDATIONS

Bid Data Analytics Architectures are going to shape the way data is received, processed, disseminated and visualized. The complexity of Big Data Analytics requirement is that specialized systems are developed to deal with enormous data exchange and data storage requirements (Charles, 2018). This research has proposed a Big Data Analytics Architecture for urban transportation systems and found out the necessary components required for data acquisition, technologies required for the data processing and an analysis of available open source architectures that can be transformed to fit into urban transportation systems.

Data is being captured in all domains of society's life, yet it is crystal clear it is not utilized. It is also clear that the pace of growth of highways does not keep up with that of vehicles on our roads. Road transportation is increasingly suffering from over congestion, leading to decreasing safety and satisfaction among road users. The future of transport lies in the disruptive nature of technology. Big Data Analytics is one of them.

5. REFERENCES

- [1] Barrachina et al., J. G. P. F. M. M. F. C. J. C. C. a. M. P., 2015. A V2I-based real-time traffic density estimation system in urban scenarios.. *Wireless Personal Communications*, pp. 259-280.
- [2] Bekiaris, L. N. R. C. a. P. M., 2017. Highway traffic state estimation per lane in the presence of connected vehicles. *Transportation research part B: methodological*. p. 106.
- [3] Charles, P. B. T. a. S. V., 2018. *BIG DATA* –

- CONCEPTS, ANALYTICS, ARCHITECTURES – OVERVIEW. *International Research Journal of Engineering and Technology (IRJET)*, February.V(2).
- [4] Columbus, L., 2016. Roundup Of Analytics, Big Data & BI Forecasts And Market Estimates. *Forbes Magazine Online*.
- [5] Daiheng, N., 2016. Traffic Sensing Technologies. *Science Direct*.
- [6] Deloitte, 2019. Global Mobile Consumer Survey, Nairobi: Deloitte Kenya.
- [7] Gokdeniz, I., 2017. Strategic Assessment based on 7S McKinsey Model for a Business by Using Analytic Network Process (ANP). *International Journal of Academic Research in Business and Social Sciences*, Volume 7.
- [8] GTFS, 2019. General Transit Feed Specification. [Online] Available at: <https://gtfs.org/>[Accessed Monday, August 05, 2019 August 2019].
- [9] Hofman, W., 2015. Data collection architecture for Big Data - a framework for a research agenda. Soesterberg, s.n.
- [10] Hoti, E., 2015. The technological, organizational and environmental framework of IS innovation adaptation in SME. Evidence from Research over the last 10 years. *International Journal of Business and Management*, III(4), p. 14.
- [11] ILO, 2017. Operations/Human Resources Manual for Matatu Saccos/Companies, s.l.: International Labour Organizations.
- [12] International Transport Forum, 2019. New Directions for Data-Driven Transport Safety, s.l.: ITF Corporate Partnership Board.
- [13] Klopp et al., J. W. S. W. P. O. D. & W. A., 2015. Leveraging Cellphones for Wayfinding and Journey Planning in Semi-formal Bus Systems: Lessons from Digital Matatus in Nairobi. In *Planning Support Systems and Smart Cities*, pp. 200-241.
- [14] KNBS , 2018. Mobile Penetration in Kenya, Nairobi: Kenya National Bureau of Statistics.
- [15] Korbil, M. S. S. K. a. N. J., 2019. Enabling a digital and analytics transformation in heavy-industry manufacturing. [Online]
- [16] Available at: Enabling a digital and analytics transformation in heavy-industry manufacturing
- [17] [Accessed 08 January 2020].
- [18] Lee, Z. e. a., 2019. Big Data Analytics in Intelligent Transportation Systems: A Survey. *IEEE Transactions on Intelligent Transportation Systems*.
- [19] NAMATA, 2017. The Nairobi Metropolitan Area Transport Authority. Orders and Acts, 17 February, 1(I), p. 75.
- [20] NIST, 2019. NIST Big Data Interoperability Framework. NIST Big Data Public Working Group Definitions and Taxonomies Subgroup, 18 March, 6(3), pp. 3-20.
- [21] Oya, Z. a. A. Ö., 2017. An Imagination of Organizations in the Future: Rethinking McKinsey's 7S Model. DOI: 10.4018/978-1-5225-1656-9.ch006.
- [22] Robert et al., B.-A. W. T. K. a. S. F., 2016. Big data analytics for transportation: Problems and prospects for its application in China. s.l., IEEE.
- [23] Sarah Williams, J. K. a. H. K., n.d. GTFS for the rest of us. Pennsylvania, s.n.
- [24] SASRA, 2019. The Sacco Societies Regulatory Authority, The Sacco Societies Act (No. 14 of 2008), Nairobi: SASRA.
- [25] Silva et al, T. & P. M. R. & B. I. & M. J. & A. D. & R. P. & N. A., 2018. Big Data Analytics Technologies and Platforms: a brief review.. Rio de Janeiro (Brazil), Conference: LADaS - Latin America Data Science Workshop.
- [26] UN Habitat., 2019. The Tools of the City Prosperity Initiative. [Online] Available at: <http://cpi.unhabitat.org/tools-city-prosperity-initiative> [Accessed Monday August 2019].
- [27] Williams S, W. A. W. P. O. D. K. J., 2016. The digital matatu project: Using cell phones to create an open source data for Nairobi's semi-formal bus system. *Journal of Transport Geography* .
- [28] Zhu et al., F. R. Y. Y. W. B. N. a. T. T., 2019. Big Data Analytics in Intelligent Transportation Systems: A Survey. in *IEEE Transactions on Intelligent Transportation Systems*, Volume Vol IV.