THE ROLE OF DATA GOVERNANCE IN THE DEVELOPMENT OF INCLUSIVE SMART CITIES

Damian Okaibedi Eke, Obas John Ebohon

De Montfort University (United Kingdom), London South Bank University (United Kingdom)

damian.eke@dmu.ac.uk; ebohono@lsbu.ac.uk

ABSTRACT

As more cities turn to information and communications technology (ICT) for the efficient delivery of public services, there are huge potentials to improve access to better infrastructure and services, including water supply and waste disposal facilities, urban transport networks, safer public spaces and improved public engagement or interaction. At the heart of this smart city initiatives is a big and robust data ecosystem that generates insights, stimulates innovation and efficiency, improves productivity and delivers wider social, economic and cultural benefits. The nature of the available data for providing these solutions, predictions and decisions can advance or impede inclusion in cities. Harnessing the benefits of smart cities for all communities is therefore mainly dependent on a functional data economy with good quality data and responsible governance approaches. The paper identifies the roles data governance can play in developing inclusive smart cities and then suggests a sustainable smart city data governance framework that is aimed at fostering inclusion by aligning with diverse objectives for and by residents.

KEYWORDS: Data governance, smart cities, inclusion, sustainability, social exclusion.

1. INTRODUCTION

In the face of the current big data economy, the idea of 'smart cities' have gained more traction in Europe and beyond. As more cities turn to information and communications technology (ICT) for the efficient delivery of public services, there are huge potentials to improve access to better infrastructure and services, including water supply and waste disposal facilities, urban transport networks, safer public spaces and improved public engagement or interaction (Pla-Castells, et al, 2014). However, smart city initiatives have been criticised for overemphasizing technological solutions and business interests over social inclusion (Paskaleva et al., 2017) - an integral part of sustainable urban development (Chan and Lee, 2008). In a 2018 discussion panel on 'The Invisible Smart city', urban designer Gil Peñalosa stated that ''we currently design our cities as though everyone is 30 and active'', indicating the exclusions of a sizeable proportion of the population outside of this energetic and active segment of the population. In a similar vein, findings of the Microsoft-backed initiative- Smart Cities for All, ''most of today's smart cities, in both the global north and the global south, are not fully accessible''. These indicate that smart city designs reflect traditional urban design biases that exclude parts of resident communities such as children, women, older population, the disabled, low income households and the

mentally ill (O'Dell et al., 2019). With about 15% of the world population living with some sort of disability (WHO, 2011), about 12.3% of the global population over the age of 60 (ONS, 2018) and with nearly half of the world's population living under the poverty line (World Bank, 2020), there is great need for prioritising inclusion in 'smart city' urban development initiatives . This is particularly important because making "cities and human settlements inclusive, safe, resilient and sustainable" where equality of access and outcome of urban opportunities is a key goal of the UN's 2030 Agenda for sustainable development. Sustainable Smart cities therefore, should advance or reinforce inclusion; appreciating the diversity of different communities, and eliminating identifiable digital, economic, physical and cultural barriers to social inclusion in urban development are critical to successful implementation of the UN SDG 2030 Goals.

At the heart of a successful smart city is a big and robust data ecosystem that generates insights, stimulates innovation and efficiency, improves productivity and delivers wider social benefits (Bibri, 2018; Hashem et al., 2016). ''Big data'' refers to the datasets that represent relevant activities that are characteristically big in volume, velocity, variety, veracity and value (Chen, et al., 2012) (Fothergill et al., 2019). Data plays a central role in the services provided in smart cities. Digital data platforms and cloud-based systems enable smart cities to collect multimodal, cross-functional, big, complex but mostly unstructured data (Chen, et al., 2014) of residents activities with associated individual and collective risks like; data protection, privacy, data sharing, environmental neglect, economic discrimination, social bias and data subject rights. Data is extracted from sources like healthcare systems, transportation, power grids, crime records, irrigation systems and other public service networks which are then used to recognize patterns and needs of the residents. While these different types of data can fuel innovation in smart cities, they can also facilitate exclusion. For instance, many of the smart city data are collected using facial recognition software but a recent study has revealed that commercial facial-recognition software show error rate of 0.8 percent for white male and 34.7 percent for black females (Buolamwini and Gebru, 2018; Raji and Buolamwini, 2019). The findings of the study demonstrate inherent racial and gender bias and further evaluation into the cause of this evident bias in the technology shows that the algorithms are informed by datasets that were lacking in diversity (Buolamwini and Gebru, 2018). This is further evidence that datasets have large influences on how technology discriminates certain groups of people in today's society.

The nature of the smart city data, its method of collection and usage have great impact on issues such as; respect for human rights, equitable distribution, respect for diversity and inclusion. While available regulations particularly the EU's General Data Protection Regulation (GDPR) focus on data subject rights, little attention is paid to the impacts of the inferential decisions made with the data which constitute exclusion of some sections of the population. The nature of the available data for such decisions can advance or impede inclusion in cities. Therefore, harnessing the benefits of smart cities for all communities is dependent on a functional data economy with good quality data and responsible governance approaches. The paper identifies the roles data governance can play in smart cities with regard to fostering inclusion. With the understanding of a smart city as ''a blend of institutions, processes, people, and technology'' (Paskaleva et al., 2017), this paper argues that an inclusive and sustainable smart city requires a sustainable data governance characterized by diverse datasets and approaches that address community, environmental, social and economic risks and concerns. The argument here is that the UN's SDG goal 11 of sustainable cities and communities cannot be achieved without a collaborative, dialogical approach to data governance where only datasets that reflect the

diverse nature of the population should be used to make inferential decisions affecting the people, the environment and the economy.

In this paper, two major questions are addressed. First, what is the relationship between data processing in smart cities and inclusion/exclusion? And second, how can data governance foster inclusion in smart cities? Answers to these questions are provided through a critical literature review. This is a non-empirical paper supported by critical review of literature on urban inclusion, smart cities and data governance. Academic literature helped us to construct an emerging narrative of smart city exclusion and the relationship with data governance. The paper is deeply rooted in analysis of the conceptual relationships between inclusion/exclusion, smart city technologies and data governance. What emerges is the overview of the roles data governance can play in a sustainable smart city that is then used to provide recommendations for an inclusive framework for data governance. We start with a detailed clarification of urban exclusion, dove-tailing into the issues of smart city exclusion exacerbated by pervasive and ubiquitous technologies. The identified roles of data governance in inclusive smart cities are then used to provide recommendations on a framework that can prioritise inclusion in smart cities. This paper offers a unique contribution to the general discourse of sustainable and inclusive smart cities. The focus on data governance illustrates the interrelatedness of data and wider social issues, particularly the inferences drawn from such data analysis. The conclusions contribute to the ever-growing discussion on the responsible data governance for AI applications. These will not only be of interest to developers of smart cities but also other experts working on AI systems and inclusion.

2. THE CONCEPT OF SOCIAL EXCLUSION AND INCLUSION

Writing about exclusion, Murard (2002 p.41) described it as an empty box given by the French state to which has since been ‘‘filled with a huge number of pages, treaties and pictures, in varying degrees academic, popular, original and valuable’’. Even though the historical roots of this concept can be traced back to ancient Greece, Murard was referring to the contemporary emergence of exclusion in France which is linked to the documented civil unrest in the late 1960s on the heels of growing unemployment and socio-economic inequalities (Ibid). It was not until the concept became prominent in national, regional and international policy agenda that attention began to shift to defining and specifying its meaning. In the late 1990s, the UK government, the European Union and the International Labour Organization helped to popularize this concept (Mathieson et al., 2008). However, a uniform definition remains difficult because the concept is largely described by different people according to its constituent elements. Wolfe, (1995) explained this by offering examples of what people can be excluded from which include; livelihood, social services, consumer culture, political choice, community solidarity and knowledge of the society and oneself. The UK Social Exclusion Unit defined it as what can happen when people or communities suffer from unemployment, poor skills, low incomes, poor housing, high crime environment, bad health and family breakdown (SEU, 1997). This definition portrays exclusion as a consequence of a number of risk factors. In the same vein, Estivill (2003) described it as a result of a combinational process that puts persons or communities at a disadvantage in relation to power, resources and prevailing values. These definitions do not provide a direct description of this concept but provide potential causes or predictors of social exclusion In 2001, the first round of the EU social inclusion process was launched and produced a Joint Inclusion Report that agreed that this concept should be defined

on the basis of a number of risk factors including but not limited to low income, unskilled labour, poor health, immigration, low education, gender inequality, discrimination and racism, age, marital status and health (Council of the European Union, 2001). These risk factors were subsequently adopted and documented as the Laeken indicators from the 2001 EU Summit in Leaken-Brussels, Belgium.

Furthermore, (Jehoel-Gijsbers and Vrooman, 2007) argued that the concept of exclusion should not only be about the process of being socially excluded but should also be about the condition of being socially excluded because exclusion can have a relational dimension (economic and structural exclusion) and a distributional dimension (socio-cultural exclusion). Both dimensions imply a lack of access to benefits owing to some socio-cultural and economic variables. From the individual level perspective, it is about any factor including race, nationality, ethnicity, age, sex, sexuality, poverty and disability that causes a person's incapacity to participate or access social activities and to build meaningful social relations (Silver, 2007). In Urban areas, this can involve insufficient integration into the social and cultural life of the society, or the insufficient access to basic needs including housing, foods, and health services amongst others. Exclusion and its opposite concept of inclusion has become a central concept in today's urban discourse.

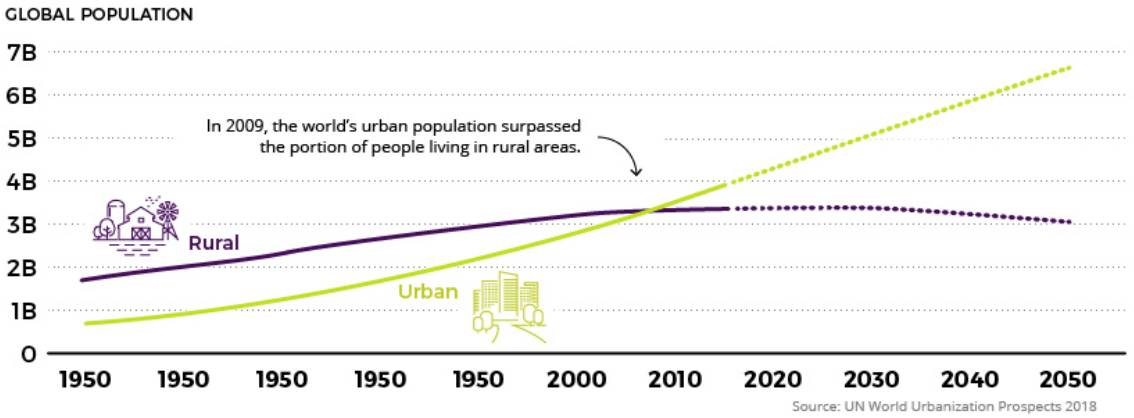
Urban cities tend to reflect the intricate social differentiations of the society, exclusion is a critical concept in urban studies and as Nowosielski (2012) argued, it is the most studied subject in the fields of urban development studies. For the purposes of this paper, we will define exclusion in urban areas as a multidimensional phenomenon where persons or communities are explicitly or implicitly denied full access or participation in the social, economic, cultural and political life of the city. It occurs when people are denied equal access to the labour market, education, healthcare, transport, judicial system, participation and other benefits in relation to others in the same city. This revolves around the concepts of poverty, marginalization and deprivation that hinders sustainable urban growth and development. Thus, myriads of initiatives are being deployed by city authorities and policy makers to tackle urban exclusions and foster inclusion to ensure that benefits of urban economies, policies and programmes benefit all sections of the society. This supports Murie and Musterd (2004) who asserted that social exclusions is a normative concept that highlights the need to address apparent inequalities or lack of participation.

Exclusion is a concept that is getting more attention considering the significant changes in contemporary cities caused by globalization and advancements in technology. Urbanization is increasing rapidly owing to increase in urban birth rates, rural-urban migration and movements of people across international borders (Serageldin, et. al., 2004).

According to Figure 1, it is observable that in 2008, the global population living in urban areas surpassed the rural population. According to United Nations Statistics which correlates with figure one above, about 55% percent of the world's population currently live in urban areas and by 2050, this is expected to reach about 68%. The dynamics of urban demographics are rapidly changing; new communities, and new industries and business demanding new set of skills and competences are emerging, resulting in new social formation and structures. It should be noted that accelerating urbanisation in many cities of the world, predominantly in the developed countries, is equally matched by cities experiencing depopulation mainly in the developed countries. Accompanying this new urban reality across global cities is the diversity and polarization, which is hardly surprising as the old economies defined by heavy industries are displaced by the new economies that are knowledge and technologically driven, requiring new

skills and competences. Such transitions in the urban economy have intensified urban inequalities that continue to manifest economically, socially, and spatially.

Figure 1. Global population growth.



The challenge for all urban development stakeholders has been how to generate inclusive urban growth and prosperity accessible to all residence irrespective of income, race, gender, sexual orientation and age; (Kearns and Paddison, 2000; Wrigley, 2002). In other words, how can the benefits of urbanization - infrastructures, services and social networks be equitably accessed to ensure inclusion. Thus, inclusion has become a popular aspirational concept in recent developmental discourse (Kasper, 2003). It is a term that has been defined by the World bank (2013: 3) as ‘‘the process of improving the terms of individuals and groups to take part in the society… improving the ability, opportunity, and dignity of people, disadvantaged on the basis of their identity, to take part in society’’. In the United Nations 2030 Sustainable Development Agenda, the concept had about 45 references as ‘inclusion’ or ‘inclusive’ and about six times in the list of sustainable development goals. It was aptly captured by the term “leaving no one behind” in SDG2. The New Urban Agenda, agreed at the Habitat III Conference also has about 45 references to the concept of inclusion (Atkinson, 2000). The aspiration reflects the desire for inclusive urban development underpinning to sustainable development. As McGranahan et al (2016) asserted, inclusion transcends more than elimination of exclusion, it involves an active process of creating equitable services and policies including proactive pursuit and guarantee of human rights. But how can this be contextualized in critical discourse of smart cities?

3. SMART CITIES

The Urban theorist and historian, Lewis Mumford, defined a city as ‘‘geographical plexus, an economic organization, an institutional process, a theater of social action, and an aesthetic symbol of creative unity’’ (Donald and Williams, 2011; Hudson, 2011). This definition portrays a city as a dynamic populous urban settlement that is essentially characterized by geographical, economic and social activities. Despite the accumulation of a significant body of literature on urban development, there is no consensus on what a city is or should be at the national and international level (Kasper et al., 2017). Kasper et al (2017) explored different concepts of this term that have emerged in academic and policy literature including; world or global cities, charter cities, prosperous cities, inclusive cities and smart cities.

Smart cities are fundamentally characterized by the use of information and communications technology (ICT), to develop, deploy, promote and improve practices and policies to address social, economic, political, cultural and environmental challenges (Caragliu et al., 2011; Leydesdorff and Deakin, 2011). It is a concept that does not enjoy a universally accepted definition. Smart city means different things to different people; with different variations in different cities depending on a lot of factors including resources and levels of aspirations, willingness to change and capacities available. However, smart city's underlying framework is that it essentially uses an intelligent network of connected objects and machines (Internet of Things - IoTs) and data to drive digital transformations to improve the lives of citizens and visitors. The World Bank offered further insights to the attributes of smart city[1] by indicating the technological intensity of smart cities, particularly the prevalence of ubiquitous sensors to access real time data. Furthermore, a smart city is projected as a city that uses technology to foster better relationships between citizens and governments. These attributes capture key elements of a smart city: technology and creating efficient relationships between key stakeholders of the city. It involves the use of data-driven technologies to connect different components of the city, optimizing public service operations and infrastructures. It is about the use of data-driven technological approaches to address challenges in waste management, public transport, policing, public health services, welfare systems, emergency response systems, infrastructure challenges and other aspects of the city life. Smart cities are driven by big, multimodal and multidimensional data from connected devices, public agencies, private companies and residents.

Data is a common element of smart cities around the world. Every smart city requires data which forms the base of the smart city model. IoT devices, sensors, networks and applications are all used to gather relevant data to enable efficiency in the technology solutions. Big data generated by IoTs and other applications are then processed and analysed to improve services and infrastructures. Big data analysis therefore plays an important role in smart city operations. From the monitoring of environmental pollution levels, wildlife counts, health monitoring of buildings and dams, traffic lights, CCTV's, connected vehicles, to smart home applications, smart cities use sensors to gather data. Smart applications are then deployed to process and analyse the data, deriving insights which inform actions such as notifications (such as parking space availability, highway incident alerts). Cisco estimates that cities that run on information/data can improve their energy efficiency by 30% within 20 years.[2] Such is the power of data in the context of smart cities. These cities are no longer aspirations but realities in Europe, the US, China and many parts of the world. Cisco, IBM, Intel, Silver Spring Networks, Build.io and Siemens are among the many tech companies providing smart city solutions covering a range of areas: hospitals, traffic/transportation, power plants, water supply and waste management. So far, global smart city spending has reached 34.35billion USD and counting.[3]

Citizens engage with smart city technologies and policies in a variety of ways involving IT skills/knowledge. For instance, some smart city infrastructures require smartphone skills such as the ability to pair mobile phones to city services. This raises the possibility of excluding some sections of the society that lack such skills or lack the financial ability to own smartphones. It does not mean that these disparities in required skills are necessarily caused by smart cities but

[1] https://www.worldbank.org/en/topic/digitaldevelopment/brief/smart-cities

[2] https://newsroom.cisco.com/press-release-content?articleId=4766225

[3] https://mobility.here.com/smart-city-technologies-role-and-applications-big-data-and-iot

the solutions they provide can exacerbate the inequalities that already exist in the society (Gilbert, 2010). Citizens are also affected by these technologies differently depending on the width and breadth of ICT applications used. An example of this is the use of Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) in the criminal justice system in some US states. COMPAS is a risk-assessment algorithm being used by counties in the US to predict crime rates, determine jail time and provide information used for sentencing. In 2016, reporters working for ProPublica analyzed about 7,000 COMPAS assessments in Broward County. The conclusion of this research was that the algorithm was biased against blacks because ''blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend''. This is a typical example of how deep learning analytics can contribute to the exclusion of residents in a smart city. These demonstrate that the data that offers powerful value to smart city initiatives can also contribute to social discrimination and exclusion. Inclusion in a smart city is therefore about the elimination/reduction of social exclusion but also about creating processes that actively protect the rights of all residents. The argument in this paper is that what and how data is used in smart cities can provide solutions to exclusion challenges in the smart city discourse. So, the question is how can data that drive smart cities be governed?

3.1. Data governance in smart cities

As it was described in the last section, data is at the core of the decision-making process of all smart cities. Multidimensional and multimodal data that include human, animal and technical data, inform smart city solutions. Like in many data-intensive operations, the multifaceted nature, volume and value of data in smart cities call for an effective data governance structure. But key questions here are: what is data governance and how can it be understood in the context of smart cities? Data governance and information governance are often used interchangeably in literature (Godinez et al., 2010) but scholars such as Kooper et al, (2011) and Nielsen (2017) stress that there are fundamental differences between the two. Kooper et al (2011) described data governance as a process that focuses on data assets while information governance is related to interactions. The data governance literature paints a picture of a term that is evolving both as a concept and as a discipline (Zhang and Yuan, 2016). Its different definitions are informed by the goal of the institution or discipline where the concept is contextualized.

Many of these definitions describe it as a compliance process to fulfil internal organizational (Donaldson and Walker, 2004) or external legal or ethical requirements (Chalcraft, 2018). This perspective focuses on responsive processes to policies, principles and regulations such as the GDPR. Others see data governance as an organizational decision-making process about data related issues (Putro et al., 2015; Weber et al., 2009). These scholars conceptualize data governance as a framework for accountability that encourages better use of data to achieve the organizations' objectives. Data governance is also defined according to how its goals are perceived. In this vein, Nielsen (2017) stated that the goal of data governance is to enhance 'business goals'. For business organizations that conceive data as key business assets, data governance becomes a way of managing assets (Aiken, 2016). This is done through critical alignment and organization of data management strategies with the business strategies (Brous et al., 2016). These definitions can all be situated in organizational literature where 'business goals' are determined by the pursuit of the bottom-line.

In a bid to offer a robust definition of data governance that offers full consideration of ethics and implementation challenges, Fothergil et al (2019) defined data governance as the overall

management of the availability, usability, integrity, quality and security of data in a given organization with the intention of ensuring maximum creation of value from the data while adhering to ethical and legal requirements. In this paper, we extend this definition to also involve the establishment of processes, policies, roles and responsibilities that foster the effective management of data for the benefit of relevant stakeholders that will be affected by the decisions derived from the data.

The EU DECODE project recently called for a different understanding of data as a common good rather than an asset in the context of smart cities (Bass, 2017). They observed that the current digital economy fosters an ecosystem where data is collected and used in ways that, 'create stark new imbalances of power' which means that 'cities will need to play a more active role in leveraging more responsible innovation with data in the local economy'. This is a call for a different kind of data governance in smart cities where data can be seen as a common good that can deliver significant personal and public benefits. We believe that one of these benefits is the building of inclusive communities where improved participation of residents can be achieved. Through a sustainable data governance framework, smart cities can achieve improved inclusion of different communities.

3.2. Sustainable data governance and Inclusion in Smart Cities

According to the European Commission, the concept of smart city means striving for sustainability through smart solutions.[4] The EC report on the Impacts of Information and Communication Technologies on Energy Efficiency[5] identified areas in a city where ICT can positively influence. Smart city is also a place where efficient ICT networks and services create benefits for all communities and for businesses through the use of cross-functional data. The underlying logic is to be more responsible, sustainable and inclusive. Achieving these sustainable, inclusive goals requires setting up a sustainable data governance structure that can enable cities to turn data into benefits for businesses, improve the quality of life of citizens (including increased access and participation), and at the same time, ensuring effective response to environmental challenges. A smart city data governance framework should focus on the people, businesses and the environment. The neglect of any of these elements can implicitly or explicitly constitute exclusion.

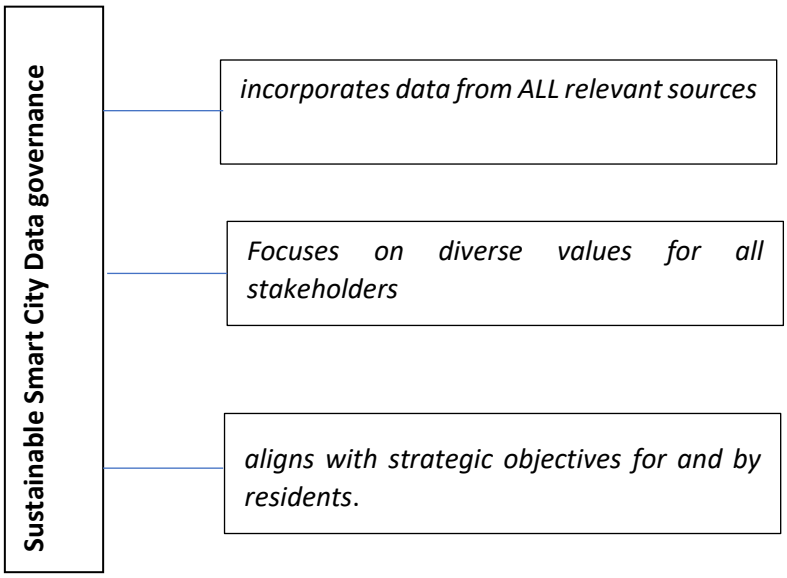
The data at the core of smart city solutions goes beyond human data to animal, technical and environmental data. Human data alone cannot provide a robust understanding of urban challenges. To harness the benefits of smart technologies (IoT, artificial Intelligence- AI) in cities, rich and diverse data from more sources provides high values. That means data governance in smart cities should not be limited to human related issues of privacy, confidentiality and data protection. It should also focus on efficiency gains and economic optimization and also on improving the environment (such as air quality, lower emissions etc). It requires a data governance framework that can help to balance competing interests of all stakeholders in the city, especially concerns of the most vulnerable residents of the city, generate economic value to support needed services and leave less environmental footprints. This is our idea of sustainable data governance which incorporates data from all relevant sources, focuses on

[4] https://ec.europa.eu/digital-single-market/en/smart-cities

[5] https://ec.europa.eu/information_society/activities/sustainable_growth/docs/studies/2008/2008_impact-of-ict_on_ee.pdf

diverse values for all stakeholders and aligns with strategic objectives for and by residents and is demonstrated in the figure below.

Figure 2. Sustainable smart city data governance.



Cities have differing needs but a city seeking sustainability and inclusiveness must focus on three key elements: socio-cultural sustainability, economic sustainability and environmental sustainability (Basiago, 1999). Socio-cultural sustainability is about how the city respects citizens' intrinsic value by ensuring that their rights to social justice, health, education, culture, religion, peace, privacy and confidentiality are protected. Economic sustainability concerns how the city provides the capacity for citizens and businesses to develop economic potentials while environmental sustainability refers to the ability of cities to protect and maintain environmental resources for future generations. Smart city data governance must focus on these three elements in order to achieve inclusivity. A narrow focus on only one or two of these areas of sustainability can exclude some communities from benefiting from smart city initiatives. A sustainable city is essentially an inclusive city. The quality, availability and integrity of the data and how they are governed have impacts on how human rights (including privacy and confidentiality) are protected, how attractive the city is to businesses, the level of government's productivity and inclusiveness, environmental sustainability and ultimately how livable the city is.

Paskaleva et al (2017) observed that this idea of harnessing smart city data to improve urban sustainability and inclusion is yet to be explored. It is something we believe is important considering the significant value data can provide and the increasing level of connectivity and collaborations between data, technology, citizens, the environment and private enterprises in smart cities. Smart city data governance goes beyond compliance and encompasses how data can be managed to create values for citizens, businesses and the environment. To derive maximum value of smart city data, sustainable data governance promotes collaboration of key partners/stakeholders in the collection of sufficient data for smart city decisions. Sustainable Development Goal is about the establishment of a resilient framework that will manage the exploding quantity of data and disparate data sources to deliver enduring value to relevant

stakeholders in the city. It is a framework that should shape the quality, availability, obtainability, usability, security and effectiveness of the data used for smart city decisions; striking the balance between data protection and acceleration of ICT potentials.

3.3. Inclusion as a priority of Sustainable data governance

Smart city data can come in many different formats, collected by multiple devices and agents in different sectors including: energy, health, government, economy, environment, community life and directly from citizens (Batty, 2013). Smart city systems will only be as good as the quality and diverse nature of the data that inform them. That is why it is necessary for smart city data to reflect the intersectional differences of the demography of the city. Non-representative datasets can lead to biased decisions exacerbating the issues of inequality and discrimination. Lack of focus on the diversity of the data that informs smart city decisions will lead to overlooking the diverse needs and preferences of the people. Aura Vasquez, a former commissioner of the Los Angeles Department of water and power observed that, ‘‘without understanding the people that are going to live in this smart city… what their priorities and problems are - we’re not going to get to them’’. It is the nature and quality of the data that will provide an improved understanding of the diverse needs of all stakeholders in the city. To be inclusive, smart cities need to use representative stakeholder data to provide critical understanding of the people and their needs.

A sustainable and inclusive smart city should also make use of diverse economic and environmental data which are traditionally collected, analysed, aggregated and utilized by different actors working towards their specific goals without appreciating the value of collaborating with others beyond their operational ‘silos’. For instance, fuel consumption records from factory machineries are required for financial accounting but can also be used for emission reporting (Gerrard, 2014). Geological information collected during mineral resource drilling for technical drilling purposes can also be used for determining waste management options and potential for acid mine drainage. These and many more environmental data have been recognized as for decision making in recognising, minimizing and environmental mitigating risk (Gerrard, 2014).

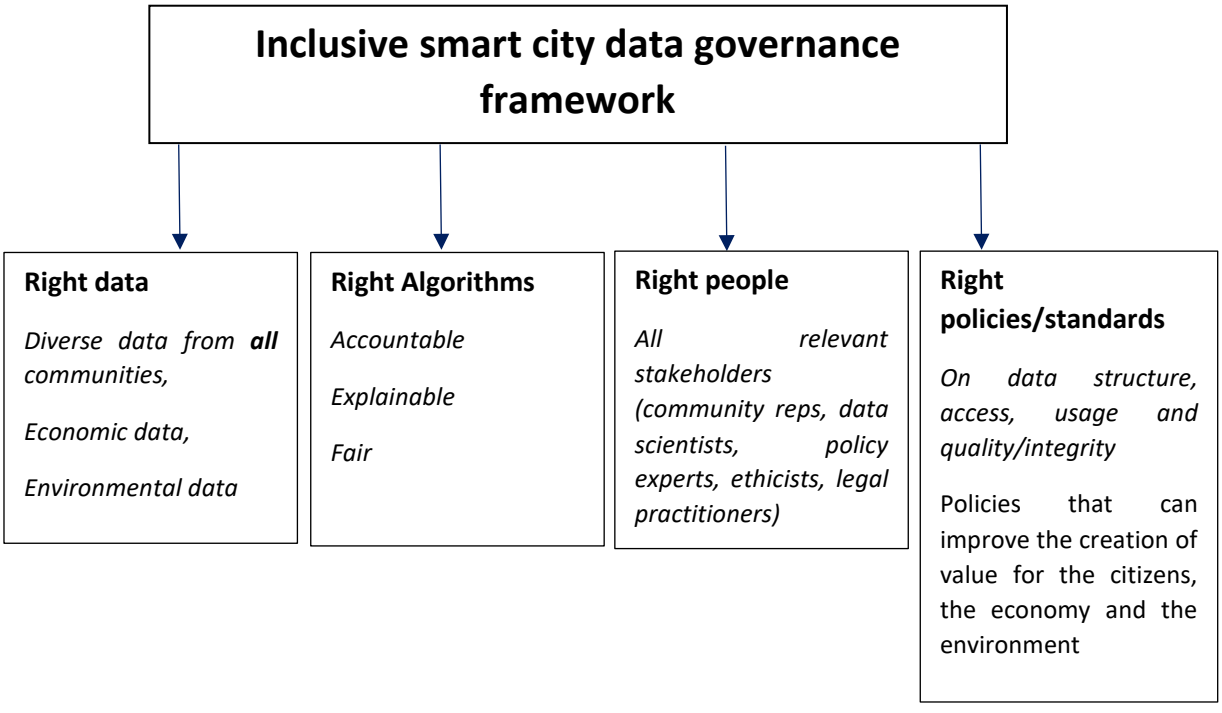
Managing the use of smart data is also the domain of data governance. Setting up a governance framework that ensures the protection of stakeholders’ human rights (including rights to privacy, to equal socio-economic and cultural participation) and builds trust should be the goal of smart cities. Sustainable data governance should proactively encourage the use of smart city data to improve the quality of life of the residents but also for economic growth and environmental sustainability. The last two are particularly important in the context of sustainable data governance. Poverty and employment status are two causal factors of social exclusion. When smart city data is used for public/private economic growth, there are possibilities generating jobs and economic opportunities that can alleviate issues of poverty and employment. Cities must develop responsible approaches to sharing relevant smart city data for business uses in ways that it will benefit the citizens. For instance, Cia et al (2018) believe that a symbiotic approach to data in a smart city can increase Cellular network operators’ efficiency and ensure lower operating costs which implicitly benefits the citizens. The physical environment can also determine how some people participate in city life. According to Williams et al., (2008), buildings design and structure, inaccessibility of public spaces and other environmental factors restrict disabled and elderly people from participating in community and

outdoor activities. Sustainable data governance of smart city data should not only focus on data protection issues but also on how to use data to ensure ‘‘equal access to resources, poverty alleviation, disaster and hazard management, land use to reduce biodiversity loss… creating a low-carbon and energy efficient society’’ (Dwevedi et al., 2018). But the practical challenge remains how to manage the diverse and big smart city data to ensure inclusive distribution of resources, fair access to services, improve mobility of all residents (including disabled people) and the livability of the city while being prepared to respond to environmental changes.

4. A FRAMEWORK FOR INCLUSION

Using the above definitions and discussion, we propose a framework for social inclusion in smart cities through data governance. This framework has no methodological or theoretical underpinnings but serves as a prescriptive framework of how the application of data can improve smart city inclusion. It is not intended to be a tool that provides solutions to all concerns of social exclusion in smart cities but rather it is presented as a structure that can help to address social, political, legal, cultural and ethical concerns that have impacts on exclusion in smart cities. Figure 3 shows should be considered in developing a data governance structure for smart cities; the right data, algorithms, people and policies/standards.

Figure 3. Framework for inclusive smart city data governance.



4.1. Right data

Data that inform smart city decisions must be right to produce inclusive outcomes. According to a recent survey of bias and fairness in machine learning, data and the algorithm are two key sources of bias in AI and machine learning (Mehrabi et al., 2019). Bias in data can be in the form of representation bias, measurement bias, evaluation bias, aggregation bias, population bias,

sampling bias or data linking bias (Ibid). It is important for smart city decisions to be based on data that reflects the diverse stakeholders in the city. When the collected data does not accurately represent the environment, the system is meant to serve, this can lead to unfair and non-inclusive decisions. The right data for inclusive smart city decisions should data that represents even the 'hard-to-reach' or 'digitally invisible' residents (O'Dell et al, 2010). Smart city data can also contain prejudice (Kallus and Zhou, 2018) or group attribution bias which can lead to non-inclusive decisions. Data that contains prejudicial views about individuals based on race, social class, nationality, gender, sexuality, educational status is not the right data for smart city decisions. Data that contains a lopsided view of a certain community is also not the right data for smart city decisions. The reliability of the insights derived from smart city data depends on the quality, nature and integrity of the data used. Smart city data governance framework must therefore include processes, partnerships and programs that can overcome trust, resource and language barriers to data collection (Stonewell et al., 2017) and ensures the mitigation of bias associated with data.

4.2. Right algorithms

In addition to using the right data, data-driven algorithms used for smart decisions should be right. According to Mehrabi et al., (2019) algorithmic bias is when the bias is not associated with the data but is added by the algorithm. Examples of biased algorithms can be found in the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) recidivism algorithm used in judicial decisions in the US. This software is more likely to assign higher risk scores to black offers than Caucasians that committed similar or more grievous offences (Rudin, Wang and Coker, 2018). In another study evaluating the fairness of algorithms for predicting juvenile recidivism, Tolan et al., (2019) discovered biases in machine learning algorithms in data-driven risk assessment. Sustainable data governance in smart cities needs to ensure that algorithms applied to smart city data are evaluated for biases. One city that has effectively done this is Johnson county Kansas where the county partnered with a private enterprise to develop a recidivist software that predicts the likelihood of re-entry into the criminal justice system (Sullivan, 2018). But rather than using the predictions derived from data from police departments and public health centres to police the residents, they use it to allocate preventive and proactive mental health resources to reduce re-offending.

4.3.Right people

Smart city data governance structure should involve the right mix of people and expertise to adequately address the issues of inclusion. Smart city governance requires a lot of companies monitoring, policy development, data quality/integrity assessment, data security and access evaluation and data management. While data scientists can bring valuable expertise in smart city data governance, social scientists, policy makers, legal practitioners and civil society groups should all be included in the data governance decision making process. The multimodal data involved in smart city decisions requires effective engagement of citizens. Therefore, the input of people who understand the issues and values associated with the associated data is highly valuable. Representatives from different communities should be represented in smart city data governance to represent the interests and goals of their communities. The shortcomings of the smart city data can only be effectively uncovered when the interests of the many different

communities are equally considered. When communities are underrepresented in governance decisions, their needs and interests may likely be neglected.

4.4. Right policies/standards

To achieve inclusiveness, data governance framework in smart cities needs to adopt policies and standards that can build trust, transparency, accountability and responsibility. The depth, size and types of data being generated by smart cities are growing exponentially by the day and many of these are strongly guarded by government or proprietary databases. The idea of limiting access to smart city data not only prevents the useful application of the data to solve common problems, it also exacerbates the lack of trust between the government and the people. Sustainable data governance policies should promote an open data platform that optimizes interoperability and ultimately contributes to the common good. One advantage of this is that it opens the door for more innovations that can restructure the economy and improve environmental sustainability. An example is using open smart city data to develop GPS-apps for the blind and visually impaired (Ryu, Kim and Li, 2014). The Barcelona City council's CityOS is based on an open-code big data platform enabling private tech company innovations boasting the economy and improving the quality of life.

Open data platforms in smart cities also contribute to improved public access to data. The citizens should be able to know what data the city holds and how they are used. This helps to build trust in government initiatives by increasing citizens participation. The citizens collectively share the ownership of the smart city data and as such should be granted a platform to know, understand and appreciate how their data is being applied. Amsterdam's Tada-data disclosed is an example of an open data initiative giving citizens control over data. The city is using active campaign and operational tools to provide clarity about data to all participating parties in the city, citizens and businesses alike. Smart city data governance policies should indeed make inclusion a strategic imperative.

5. CRITICAL DISCUSSION AND CONCLUSION

Observably, cities play critical roles in economic development of nations, this is evidenced by the disproportionate share cities account in total gross domestic product of nation states. Nonetheless, cities also account for significant proportion of factors responsible for environmental degradation hence the desire to develop cities into sustainable and livable communities. This translates to creating employment and income generating opportunities, safe and affordable housing, and building resilient economies and societies, as we have seen emerging in many countries, particularly the developed countries. Other manifestations of sustainability in cities include investment in sustainable public transport and public spaces to afford recreational facilities for healthy living and wellbeing.

However, the challenge has been how to ensure that the opportunities sustainable cities have to offer is equitably accessed by all citizens, and many cities have mainstreamed citizens' participation into urban planning and management to bridge the widening inequality in cities. For effective participation and delivery of adequate and right amount of services in real time, the use of information technology has become inevitable. Evidence abound showing that huge amount of data is required to determine the kinds of urban services demanded by citizens and also to determine the efficiency of such deliveries in terms of indications of who are able to

access services and those who might have been left behind. The governance of the huge data input to smart cities has become just as important as the inequality the data set serves to bridge. For inclusiveness, data governance framework is a necessary prerequisite in smart cities to effect sustainable and livable urban environment. Wider access to smart city data is the only way equality of access and outcome of the opportunities cities have to offer its citizens, allowing efficient and effective delivery of urban services. Thus, data governance is critical to smart cities and the delivery of 21st century sustainable and livable cities.

REFERENCES

Aiken, P. (2016) Experience: Succeeding at data management—BigCo attempts to leverage data. *Journal of Data and Information Quality (JDIQ)*, 7(1–2), pp. 1–35.

Atkinson, R. (2000) Combating Social Exclusion in Europe: The New Urban Policy Challenge. Urban Studies, Vol. 37, No. 5-6, 1037-1055

Basiago, A. D. (1999). Economic, social, and environmental sustainability in development theory and urban planning practice

Bass, T. (2017) *Reclaiming personal data for the common good*. [Online] https://decodeproject.eu/blog/reclaiming-personal-data-common-good.

Batty, M. (2013) Big data, smart cities and city planning. *Dialogues in Human Geography*, 3(3), pp. 274–279.

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, pp. 230–253.

Brous, P., Janssen, M. and Vilminko-Heikkinen, R. (2016) Coordinating decision-making in data management activities: a systematic review of data governance principles. In: *International Conference on Electronic Government*. Springer, pp. 115–125.

Buolamwini, J. and Gebru, T. (2018) Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In: *Proceedings of Machine Learning Research*. Conference on Fairness, Accountability, and Transparency. pp. 1–15.

Caragliu, A., Del Bo, C. and Nijkamp, P. (2011) Smart cities in Europe. *Journal of urban technology*, 18(2), pp. 65–82.

Chalcraft, J. (2018) Drawing ethical boundaries for data analytics. *Information Management*, 52(1), pp. 18–25.

Chan, E., Lee, G.K.L. Critical factors for improving social sustainability of urban renewal projects. Soc Indic Res 85, 243–256 (2008). https://doi.org/10.1007/s11205-007-9089-3

Chen, H., Chiang, R.H.L. and Storey, V.C. (2012) Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36, pp. 1165–1188.

Chen, M., Mao, S. and Liu, Y. (2014) Big Data: A Survey. *Mobile Networks and Applications*, 19(2), pp. 171–209.

Council of the European Union (2001) *Joint Report on Social Inclusion*. European Commission.

Dalla Cia, M. et al. (2018) Using Smart City Data in 5G Self-Organizing Networks. *IEEE Internet of Things Journal*, 5(2), pp. 645–654.

Donaldson, A. and Walker, P. (2004) Information governance—a view from the NHS. *International journal of Medical informatics*, 73(3), pp. 281–284.

Donald, A., & Williams, A. (2011). Introduction: The Paradoxical City. In The Lure of the City: From Slums to Suburbs (pp. 1-11). London: Pluto Press. doi:10.2307/j.ctt183pc6x.3

Dwevedi, R., Krisha, V. and Aniket Kumar (2018) Environment and Big Data: Role in smart cities of India. *Resources*, 7(4), p. 64.

Estivill, J. (2003) *Concepts and strategies for combating social exclusion: an overview*. International Labour Organization.

Fothergill, B.T. et al. (2019) Responsible Data Governance of Neuroscience Big Data. *Frontiers in Neuroinformatics*, 13, p. 28.

Friedmann, J. and Wolff, G. (2017) World city formation: an agenda for research and action. In: *The Globalizing Cities Reader*. Routledge, pp. 46–54.

Gilbert, M (2010) THEORIZING DIGITAL AND URBAN INEQUALITIES, Information, Communication & Society, 13:7, 1000-1018, DOI: 10.1080/1369118X.2010.499954

Hudson, A. (2011). The Dynamic City: Citizens Make Cities. In Williams A. & Donald A. (Authors), The Lure of the City: From Slums to Suburbs (pp. 12-31). London: Pluto Press. doi:10.2307/j.ctt183pc6x.4

Godinez, M. et al. (2010) *The art of enterprise information architecture: a systems-based approach for unlocking business insight*. Pearson Education.

Hashem, I.A.T. et al. (2016) The role of big data in smart city. *International Journal of Information Management*, 36(5), pp. 748–758.

Jehoel-Gijsbers, G. and Vrooman, C. (2007) Explaining social exclusion.

Kallus, N. and Zhou, A. (2018) Residual Unfairness in Fair Machine Learning from Prejudiced Data. *arXiv:1806.02887 [cs, stat]*, [Online] Available from: http://arxiv.org/abs/1806.02887 [Accessed 12/03/2020].

Kasper, E. et al. (2017) *Inclusive Urbanisation and Cities in the Twenty-First Century*. IDS.

Kearns, A. and Paddison, R. (2000). New Challenges for Urban Governance. Urban Studies, Vol. 37, No. 5-6, 845-850.

Kooper, M.N., Maes, R. and Lindgreen, E.R. (2011) On the governance of information: Introducing a new concept of governance to support the management of information. *International journal of information management*, 31(3), pp. 195–200.

Leydesdorff, L. and Deakin, M. (2011) The triple-helix model of smart cities: A neo-evolutionary perspective. *Journal of urban technology*, 18(2), pp. 53–63.

Mathieson, J. et al. (2008) Social Exclusion Meaning, measurement and experience and links to health inequalities. *A review of literature. WHO Social Exclusion Knowledge Network Background Paper*, 1, p. 91.

McGranahan, G., Schensul, D. and Singh, G. (2016) Inclusive urbanization: Can the 2030 Agenda be delivered without it? *Environment and Urbanization*, 28(1), pp. 13–34.

Mehrabi, N. et al. (2019) A survey on bias and fairness in machine learning. *arXiv preprint arXiv:1908.09635*.

Murard, N. (2002) Guilty victims: social exclusion in contemporary France. *Biography and social exclusion in Europe. Experiences and life journeys*, pp. 41–61.

Murie, A. and Musterd, S. (2004) Social exclusion and opportunity structures in European cities and neighbourhoods. *Urban studies*, 41(8), pp. 1441–1459.

Nielsen, O.B. (2017) A comprehensive review of data governance literature. *Selected Papers IRIS*, 8, pp. 120–133.

Nowosielski, M. (2012) Challenging Urban Exclusion? Theory and Practice. *Polish Sociological Review*, 179(3), pp. 369–383.

O'Dell, K. et al. (2019) *Inclusive Smart Cities: Discovering Digital solutions for all*. Deloitte Center for Government Insights.

O'Dell, L. et al. (2010) Constructing 'normal childhoods': Young people talk about young carers. *Disability & society*, 25(6), pp. 643–655.

ONS (2018) *Living longer: how our population is changing and why it matters*. [Online] https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/articles/livinglongerhowourpopulationischangingandwhyitmatters/2018-08-13.

Paskaleva, K. et al. (2017) Data Governance in the Sustainable Smart City. *Informatics*, 4(4), p. 41.

Pla-Castells, Marta & Martinez-Durá, Juan & Samper Zapater, J. Javier & Cirilo Gimeno, Ramon. (2015). Use of ICT in Smart Cities. A practical case applied to traffic management in the city of Valencia. Conference: Smart Cities Symposium Prague 2015, Volume: 978-1-4673-6727-1 EEE10.1109/SCSP.2015.7181559.

Putro, B.L., Surendro, K. and Herbert (2016) Leadership and culture of data governance for the achievement of higher education goals (Case study: Indonesia University of Education). In: *AIP Conference Proceedings*. AIP Publishing LLC, p. 050002.

Raji, I.D. and Buolamwini, J. (2019) Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society - AIES '19*. the 2019 AAAI/ACM Conference. Honolulu, HI, USA: ACM Press, pp. 429–435.

Romer, P. (2010) *Technologies, rules, and progress: The case for charter cities*.

Rudin, C., Wang, C. and Coker, B. (2018) The age of secrecy and unfairness in recidivism prediction. *arXiv preprint arXiv:1811.00731*.

Ryu, H.-G., Kim, T. and Li, K.-J. (2014) Indoor navigation map for visually impaired people. In: *Proceedings of the Sixth ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*. pp. 32–35.

Sassen, S. (2013) *Deciphering the global: its scales, spaces and subjects*. Routledge.

Serageldin, M., Cabannes, Y., Solloso, E,. Valenzuela, L. (2004). Migratory Flows, Poverty and Social Inclusion in Latin America. World Bank Urban Research Symposium. https://www.researchgate.net/profile/Yves_Cabannes/publication/32893681_Migrator

y_Flows_Poverty_and_Social_Inclusion_in_Latin_America_By/links/55a3e9db08ae5e82ab1f25c6/Migratory-Flows-Poverty-and-Social-Inclusion-in-Latin-America-By.pdf

SEU (1997) Social Exclusion Unit: Purpose, work priorities and working methods. *London: The Stationery Office*.

Silver, H. (2007) The process of social exclusion: the dynamics of an evolving concept. *Chronic Poverty Research Centre Working Paper*, (95).

Stonewall, J. et al. (2017) Best practices for engaging underserved populations. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, pp. 130–134.

Sullivan, R. (2018) *Innovations in Identifying people who frequently use criminal justice and healthcare systems*. [Online] https://www.prainc.com/innovations-identification-cj-healthcare/.

Tolan, S. et al. (2019) Why machine learning may lead to unfairness: Evidence from risk assessment for juvenile justice in catalonia. In: *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*. pp. 83–92.

Weber, K., Otto, B. and Österle, H. (2009) One size does not fit all—a contingency approach to data governance. *Journal of Data and Information Quality (JDIQ)*, 1(1), pp. 1–27.

WHO (2011) *World report on disability 2011*. World Health Organization.

Williams, B. et al. (2008) Experiences and expectations of disabled people. *London: Office for Disability Issues*.

Wolfe, M. (1995) Globalization and social exclusion: Some paradoxes. *Social Exclusion: Rhetoric Reality Responses/International Institute for labour studies. United Nations development program*, pp. 81–102.

World Bank (2020) *Poverty*. [Online] https://www.worldbank.org/en/topic/poverty/overview. Available from : https://www.worldbank.org/en/topic/poverty/overview.

WorldBank (2013) *Social Inclusion*. [Online] https://www.worldbank.org/en/topic/social-inclusion#4.

Wrigley, N., Guy, C., Lowe, M. (2002). Urban Regeneration, Social Inclusion and Large Store Development: The Seacroft Development in Context. Urban Studies, Vol. 39, No. 11, 2101–2114.