AI Regulation for Smart Cities: Challenges and Principles

Ya Zhou and Atreyi Kankanhalli

Abstract To reduce the burden on city resources and improve governance and services, artificial intelligence (AI) is being applied toward building smart cities. Although smart city AI applications bring benefits of automation and efficiency, they also raise regulatory challenges, with concerns about discrimination in service provision, privacy, legal, and ethical issues. To address these issues, we propose an integrated and systematic framework for smart city AI regulation. To develop the framework, we first identify the dimensions of a smart city and the components of AI-enabled systems. We then discuss examples of smart city AI applications and the regulatory challenges they pose. Last, we describe our integrated framework of principles for smart city AI regulation, which addresses these challenges. Our study contributes to both the literatures on smart cities and AI frameworks, as well as practice on smart city AI policy and regulation.

Keywords Smart city artificial intelligence regulation · Artifical intelligence regulation · Smart principles · Artificial intelligence policy

Introduction

Due to urbanization and the increase in city density, the burden on city resources and governance is intensifying. To alleviate such burden as well as improve the quality of life of citizens, governments are investing in advanced information and communication technologies (ICT) and applying them to various city services, which facilitates the development of "smart cities" (Jucevičius, Patašienė, & Patašius, 2014). An emerging trend in the use of ICT is the application of artificial

Y. Zhou (✉)
Xiamen University, Xiamen, China
e-mail: zhouya@xmu.edu.cn

A. Kankanhalli
National University of Singapore, Singapore, Singapore
e-mail: atreyi@comp.nus.edu.sg

intelligence (AI), which has experienced a resurgence in recent years due to advances in techniques, such as natural language processing, machine learning, computer vision, and robotics. This is also aided by the increase in the amount of data available, i.e., big data, and the dramatic improvement of computing power (Russell & Norvig, 2016; Stoica et al., 2017).

These AI applications bring opportunities of automation and efficiency for smart city infrastructure and services, such as autonomous vehicles for transport, and detection and tracking of criminals for public safety. However, there is also the potential for negative implications during the development and application of AI for smart cities, such as discrimination in service provision, privacy, legal, and ethical issues (Stoica et al., 2017; Stoyanovich, Abiteboul, & Miklau, 2016). Hence it is crucial for governments to regulate the design and use of AI for smart city applications.

To address the regulatory challenges posed by AI systems and applications, government agencies, companies, and research institutes have proposed several frameworks to guide the development and use of AI (see Appendix). Among these frameworks (e.g., Accenture, 2016; European Commission, 2018; Microsoft, 2019; UK Government, 2019), the most common principles are transparency, equality or non-discrimination, accountability, and human values. Data privacy and safety principles have also gained attention in many of these frameworks (e.g. Japanese Society for AI, 2017; Microsoft, 2019). A few frameworks mention the need for responsible and sustainable use of AI systems (e.g., Accenture, 2016; PwC, 2019). However, existing frameworks generally propose guiding principles without a systematic, conceptual derivation, and integration of the principles. Additionally, the frameworks do not discuss the enforcement mechanisms to operationalize these principles. Furthermore, the principles are typically not adapted to the context of smart cities and thereby not linked to the dimensions, applications, and challenges of smart cities.

Motivated thus, we aim to understand and address these issues by developing a framework of principles for AI regulation for smart cities in a conceptual and systematic way. This can help researchers and practitioners to understand the principles for governing smart city AI applications and examine or develop policies and regulations accordingly. Specifically, our framework is built by addressing the following research questions: (1) What are the dimensions of a smart city? (2) What are the components of AI systems? (3) What are the regulatory challenges posed by smart city AI applications? (4) What principles should be followed to regulate smart city AI applications to address these challenges?

Thereby, we propose such an integrative framework for principles of smart city AI regulation, which covers input, algorithm, and outcomes of AI systems. Our framework contributes to both research and practice. Specifically, the framework adds to the literatures on smart city and AI frameworks by conceptually integrating principles of smart city AI regulation based on the identification of smart city dimensions and AI system components. Our framework proposes these principles to address the specific regulatory challenges posed by smart city AI applications in the real world. Furthermore, we discuss the enforcement mechanisms, such as regula-

tions, for operationalizing these principles, which is overlooked in existing AI frameworks. Finally, the framework provides practitioners i.e., policy makers and companies, with a systematic set of regulation principles for governing and developing smart city AI applications.

Before we introduce the framework, we first delve into the dimensions of a smart city and understand the components of AI systems. Then we introduce several examples of smart city AI applications and discuss the regulatory challenges entailed in these applications. Finally, we propose an integrative framework taking into account smart city dimensions and AI system components, which outlines the principles for regulation of smart city AI applications.

Dimensions of a Smart City

Smart city has been referred to as an urban ecosystem that places emphasis on the use of digital technology, shared knowledge, and cohesive processes to underpin citizen benefits (Juniper Research, 2018). However, it is hard to provide a precise definition of smart city (Jucevičius et al., 2014). Nevertheless, we could deepen our understanding of the concept through investigating its dimensions. Figure 1 shows the overview of the proposed dimensions. We derive these dimensions based on the perspective of the enabling resources for smart cities, as well as the two key aspects of smart city functioning, i.e., city governance and urban life.

Specifically, a smart city is comprised of several dimensions, which could be discussed in terms of the *city functioning* and the *enabling resources*. On one hand, smart city implies the application of smart computing technologies in the subsystems of the city. For example, Dirks and Keeling (2009) identify nine city systems, i.e., transportation, energy, education, healthcare, buildings, physical infrastructure, food, water, and public safety. According to the aspects of urban life, Lombardi, Giordano,
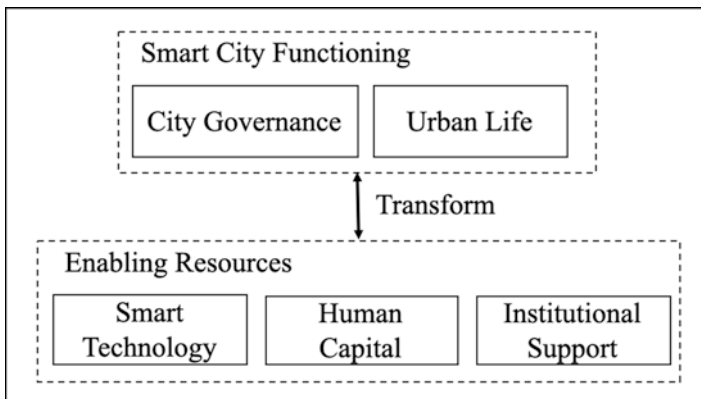

Fig. 1 Dimensions of a smart city

Farouh, and Yousef (2012) summarized six components of a smart city, i.e., smart economy for industry, smart people for education, smart governance for e-democracy, smart mobility for logistics and infrastructures, smart environment for efficiency and sustainability, and smart living for security and quality. Particularly, the former classification focuses on the *city governance* perspective, which emphasizes the management of important resources to run a smart city, while the latter views smartness from an *urban life* perspective, which highlights key areas for the growth of the smart city and the improvement of citizens' life. The further development of a smart city could go beyond the independent use of technologies in these domains and link them as an organic system to support city sustainability and better life of citizens (Albino, Berardi, & Dangelico, 2015; Nam & Pardo, 2011). Particularly, due to the interdependencies across smart city systems, joint advances in the technologies for different domains is needed. For instance, with the interconnected smart city functioning, a smart city could garner information regarding its physical infrastructure and use it to enable multiple conveniences such as facilitate mobility, conserve energy, and improve the quality of air and water (Nam & Pardo, 2011).

At the same time, a smart city development is driven by various *enabling resources* in combination (Nam & Pardo, 2011). First, *smart technology* refers to underlying intelligent IT infrastructure, systems, and applications. This includes systems with real-time awareness of the physical world and capabilities of advanced analytics to improve city infrastructure and services. Specifically, city processes can now be monitored by various approaches in real-time, such as through connected telecommunication networks, digitally controlled utility services and infrastructure, sensor and camera networks, and mobile technologies used by citizens (Kitchin, 2014). These approaches generate vast amounts of data that can be used to represent, model, and predict urban processes and simulate the likely outcomes of future urban development. With the radical improvement in computing technologies and capabilities, advanced data analytics and AI have been introduced into various city sectors to support decision-making and service provision. The combination of big data and ubiquitous intelligence together help enable a city to run in an intelligent and coordinated way (Girtelschmid, Steinbauer, Kumar, Fensel, & Kotsis, 2014).

In addition to technology, human-related resources, such as education, innovation, and learning, are also important to drive a smart city. Such *human capital* contributes to the collective intelligence and leads to a creative environment for a smart city. For example, through promoting IT education and initiating life-long learning programs, a city can not only equip its citizens with IT capabilities to utilize the services of a smart city, but also build intelligence for creating smart city innovations (Albino et al., 2015; Nam & Pardo, 2011). *Institutional support* refers to the support from government agencies and the community for governance of a smart city. The transformation through technology interacts with the political environment of the city. To transform existing cities into smart cities, government agencies not only need to provide policy and administrative support (e.g., initiatives, structure), but also to engage diverse stakeholders (e.g., citizens and companies) into the governance and facilitate the collaboration across different departments and communities (Albino et al., 2015; Nam & Pardo, 2011).

It is worth noting that smart city functioning is not only transformed by enabling resources, but is also a transformer of resources. This follows Giddens' structuration theory, which discusses the mutual interactions between technology and humans (Giddens, 1984). In other words, the development of smart city functioning could in turn facilitate the availability of enabling resources. For example, if they witness progress in city governance and improvement in urban life through applying intelligent technologies, governments, and communities would likely invest more resources into developing the technologies and thereby further promote the smart transformation of city functioning. In this sense, advances in smart city functioning could also enhance institutional support. Additionally, smart education, which is an important smart city function, could strengthen human capital by enabling more citizens to access and acquire advanced IT knowledge and resources through various channels. This would contribute to the collaborative environment for the further development of the smart city. Last, the advances in smart city functioning could lead to greater communication and interactions among different city functions, which would bring new research and development opportunities for building smart technologies.

In sum, a smart city development not only entails the adoption of intelligent computing technologies in citizen life and city governance, but also depends on human capital and the support from institutions. This in turn stimulates the production of these enabling resources.

Components of AI Systems

With regards to smart technology, AI refers to the machines or software that attempt to think and behave in a human and rational manner (Russell & Norvig, 2016). The aim of AI not only includes the successful fidelity to human performance, but also the achievement of rational or ideal performance. AI has largely been applied to reasoning, problem solving, knowledge representation, natural language processing, perceptions (e.g., face recognition, speech recognition, computer vision) and motion manipulation (e.g., robotics) (Flasiński, 2016). Figure 2 shows the components of AI systems. Particularly, we derive this figure based on an input—process—output—outcome systems view.

Similar to humans, AI systems need to capture and *input data* as the observations of the environment to make decisions and solve problems. In the past two decades, the increasing popularity of online services, such as Google Search, Amazon, and YouTube, are generating huge amounts of data in various forms, such as text, picture, video, and transactions, which can serve as inputs for AI systems. Moreover, the widespread adoption of sensors and IoT in different domains is further propelling the use of big data, where continuous data points and their network could be captured to reflect a dynamic and real-time environment. One such example is a current urban system, where different data sources are connected, such as fixed and wireless telecom networks, digitally controlled utility services, and transport infra-
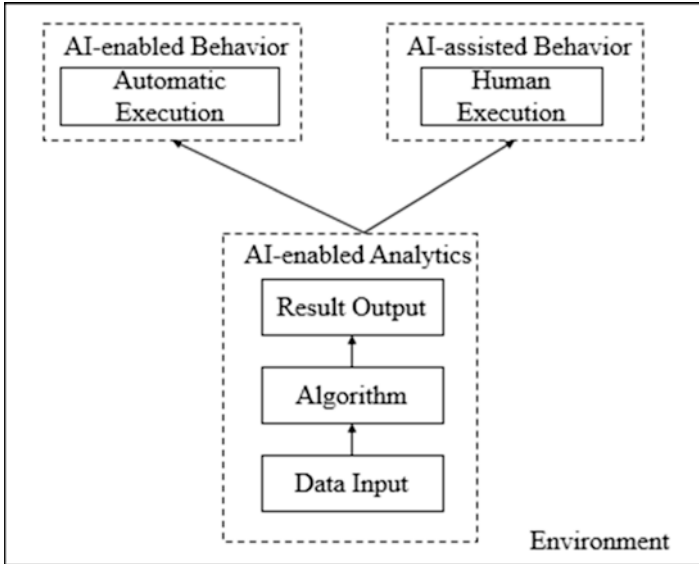
Fig. 2 Components of AI systems

structure, sensor and camera networks, building management systems, and so on. These data sources are widely used to monitor, manage, and regulate city flows and processes (Kitchin, 2014). For instance, networked cameras on roads and in public places could provide data input for AI face recognition technologies to detect and track criminals for enhancing public safety (Greene, 2018).

According to their *algorithms*, AI systems would process these inputs and generate outputs for decision-making or problem solving. Among AI algorithms, machine learning techniques have become prominent in recent years and have led to the resurgence of AI. Machine learning techniques can be divided into two types, i.e., supervised and unsupervised learning. Supervised learning involves a labeled training set, which helps systems to understand new data based on existing labels or categories. It is typically used for classification tasks. Nowadays, it is possible to acquire vast amounts of labeled data by recruiting online workers on sites such as Amazon Mechanical Turk to carry out labeling tasks, or from large online datasets, e.g., Google Open Images Dataset with more than nine million images and YouTube-8M with more than seven million labeled videos (Heath, 2018). Based on labeled data, algorithms learn how to apply these labels to new data. For example, based on previous patterns of various road obstacles, the algorithm can learn how to identify different types of road obstacles in new images or videos (Brisimi, Cassandras, Osgood, Paschalidis, & Zhang, 2016).

On the other hand, unsupervised learning identifies patterns and structures in an unlabeled data set. It is usually applied in clustering tasks. Traditional classification and regression algorithms, such as support vector machines, naïve Bayes and decision trees, fall under supervised learning, while K-means is unsupervised. For

example, combining GIS information and crime data, high and low spots of crimes can be clustered such that people can easily identify locations with high or low crime rates (GIS Geography, 2018). As a subset of machine learning techniques, neural networks and deep learning methods include multiple algorithms, which could work in either a supervised or unsupervised way. Indeed, the advances in neural network and deep learning algorithms have greatly enhanced the application of AI systems to solving problems in the real world.

According to the algorithm, the system processes inputs and then generates *outputs*. The output can be in the form of information or knowledge for decision-making or problem solving, or even instructions for robotics. Based on the outputs, actions can be taken in the real world automatically (*AI-enabled*), e.g., through robots or other automatic machines, or semi-automatically (*AI-assisted*), e.g., executed by humans who use the results as references for decision-making.

Smart City AI Applications and Regulatory Challenges

Working with ubiquitous data sources provided by the Internet of Things (IoT) as well as networked databases, AI brings opportunities of automation and efficiency for smart city systems. Indeed, AI applications are being implemented or will be implemented in various aspects of a smart city (Tech Wire Asia, 2018). However, large-scale deployment of AI technologies e.g., data-driven predictive systems, may result in unintended adverse societal consequences. To illustrate the issues, we discuss smart city AI applications and regulatory challenges mainly from the US, EU, and Singapore, which is considered as a leader in AI readiness.[1]

For example, taking advantage of AI and sensors, autonomous vehicles are becoming capable of sensing their environment and navigating without human intervention (Urban Redevelopment Authority of Singapore, 2017). Developing and deploying autonomous vehicles to public transport services can reduce the burden on manpower (e.g., bus or train drivers). Further, these vehicles may help save lives overall by more strictly conforming to transportation rules than human drivers (Chessen, 2017). However, the convenience and potential brought by the automation in mobility may also lead to ethical dilemmas. One example is the trolley problem. Autonomous self-driving vehicles would need to figure out how to respond in situations where collisions are unavoidable. In such cases, the algorithm would have to decide how to minimize harm, which could pose an ethical dilemma between options such as minimizing the number of deaths or saving the driver and passengers (Nyholm & Smids, 2016). In addition to such ethical problems, autonomous self-driving cars may also lead to legal issues. One instance is the liability of accidents. When a human driver is not present, the question arises as to whether the

---

[1] https://www.oxfordinsights.com/ai-readiness2019.

damage caused in an accident can be attributed to the owner of the car, or whether only the manufacturer can be held accountable (Chessen, 2017).

Other autonomous machines, such as healthcare robots, share similar issues. AI-enabled robots are being developed and implemented in the departments of hospitals for tasks, such as surgery, rehabilitation, and assistance (Healthcare Asia, 2017; Khalik, 2015). Additionally, they can be used for home care of patients and the elderly (Kidal, 2017). However, though these robotics technologies can improve healthcare efficiency and reduce the burden on healthcare manpower, they may pose serious concerns for patient safety, autonomy, control, and accountability. For example, according to the US FDA, there were 144 deaths and more than 1000 injuries that involved robotic surgeons between 2000 and 2014 (Thomson, 2015). Of these, 60% accidents were caused by device malfunctions, such as out of power, incorrect movement, electrical sparks, and falling pieces (MIT Technology Review, 2015). Further, if the command to a robot would harm patient safety, such as a senior citizen requesting the robot to kill him or her, or the doctor requesting a wrong medical procedure, should the robot still carry out that command? Such scenarios raise questions about who will be responsible for the adverse outcomes of these healthcare robots (Simpson, 2016).

In addition to enabling automation, AI has also been applied to support the decision-making of various smart city functions. One such important area is public safety. The networked cameras from roads and public places provide data input for face recognition technologies to track criminals (Greene, 2018). This data source can also be connected with other data sources, such as social media, Internet use, hotel stay and trips to monitor criminals' tracks. However, the ubiquitous surveillance could make people feel vulnerable about their rights to privacy, confidentiality, and freedom of expression (Gasser & Almeida, 2017). Other serious legal and societal consequences could arise when AI is used for predictive policing crime prevention. For example, there could be racial bias due to a biased training set or algorithms in the AI system. A notable case is the COMPAS predictive policing system (Angwin, Larson, Mattu, & Kirchner, 2016). This algorithm aimed to predict recidivism in individuals detained by the police, by assigning them a 'risk score'. The algorithm was found to be heavily biased against persons of color i.e., it would consistently assign a higher risk score to people of color (even though those individuals were later found to be relatively low risk), and lower scores to white persons (though they turned out to be much likelier to commit crimes later on).

Bias in AI systems could also cause discrimination in other smart city domains. For example, by applying AI in logistics system, companies attempt to optimize their arrangement of materials and product delivery. However, reporters found that Amazon's same-day delivery service was unavailable for ZIP codes in predominantly black neighborhoods (Crawford, 2016). Further, through use of AI, businesses and advertisers have opportunities to precisely reach a target population through search engines, such as Google. However, it was found that women were less likely than men to be shown ads on Google for highly paid jobs. This biased output could be caused by the biased training dataset used for job search or the algorithm in the AI system.

The unforeseeable societal outcomes of these AI applications can be largely attributed to a key characteristic: they are typically black-box learning systems. This means that their underlying decision-making and learning mechanisms are not observable by external auditing and law enforcement agencies. What's worse, even those who design these systems often do not fully understand how certain decisions are made. State-of-the-art learning algorithms often favor precision over transparency. Indeed, in many cases there is an inherent tradeoff between transparency (how well can the system be understood by a stakeholder) and accuracy (how well can the system predict the outcomes of future inputs). Ensuring transparency in AI and machine learning has seen increased attention from government bodies (Goodman & Flaxman, 2016; Smith, 2016), legal experts (Roggensack & Abrahamson, 2016), and the media (Hofman, Sharma, & Watts, 2017; Smith, 2016).

Further, in various scenarios, activities of individual users e.g., their purchases and preferences, health data, online and offline transactions, photos they take, commands they speak into their mobile phones, locations they travel to, are used as training data. This raises issues of privacy and security regarding individual's data collected, stored, processed, shared, and output from such systems. Particularly, as AI systems typically involve cloud computing and networks for data storage, communication, and analytics, cyber threats could have serious impacts on these systems. For example, in 2017, over 99 billion records were exposed due to data breaches in cloud services (Balakrishnan, 2018).

Framework of Principles for Smart City AI Regulation

In order to resolve these issues, this chapter develops an integrative framework, which can help researchers and practitioners to understand the strategies and principles for regulating AI applications for smart cities. As discussed in the previous section (see Fig. 2), AI systems include analytics and AI-enabled or assisted behaviors. Figure 3 shows the general principles for both the *analytics* and the outcome *behaviors*.

As can be seen in Fig. 3, to support the smart city (see Fig. 1), certain principles need to be implemented by regulating both AI-based behaviors and analytics design, such as protection of human rights, ethics and fairness (Gasser & Almeida, 2017; Stoica et al., 2017). Traditionally, regulations and laws have been widely implemented to protect various *human rights*. For example, everyone shall be entitled the right to life, and no one shall be arbitrarily deprived of his or her life. In principle, AI-enabled or assisted behaviors should also not hurt human beings and violate their human rights (Etzioni, 2017). However, as existing laws and rules focus on human behaviors, adaptations need to be made to regulate AI-enabled or assisted behaviors. For example, in order to regulate self-driving car trials, the Singapore government has amended the Road Traffic Act to recognize that a motor vehicle does not necessarily have a human driver. This can exempt autonomous vehicles and their operators from existing legislation set for human driving behaviors. Instead, the operators are asked to ensure there is liability insurance, or place a
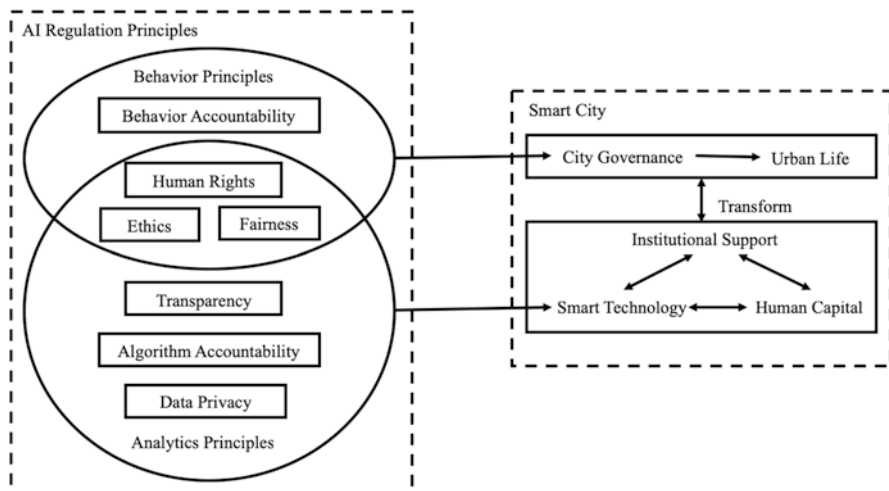
Fig. 3 Framework of principles for smart city AI regulation

security deposit (Straits Times, 2017). By explicitly defining the responsibilities of different parties in protecting human rights, this could also contribute to *behavior accountability*, which includes the clarification of liability of AI-enabled or assisted outcomes. In addition to the protection of basic human rights, the AI-enabled or assisted behaviors should also be ethical and fair (Stoica et al., 2017). For example, to alleviate concerns over bias in policing, California State has introduced a bill called the Body Cameral Accountability Act, which seeks to prohibit the use of facial recognition in police body cameras. The state has also introduced another bill to require businesses to publicly disclose their use of facial recognition technology in view of public concerns (Nonnecke & Newman, 2019).

Further, as the output of AI systems determines or influences the corresponding outcomes, conformity to these core principles of behaviors, i.e., protection of human rights, *ethics, and fairness,* would largely depend on the regulation of the AI system design. To achieve this, the design of AI systems should avoid biased data and algorithms, build rules to protect human rights, and take ethics into consideration. For example, by identifying the bias of data and algorithms in smart city AI cases, the Global Future Council on Human Rights published a white paper to propose several principles for avoiding discrimination in AI systems, such as active inclusion of the designers with diverse background, clear definition of fairness to guide system development, and visible avenues for redress for those affected by disparate impacts (Global Future Council on Human Rights, 2018). Another example is the Federal Automated Vehicles Policy in the U.S., which provides guidelines for assessing automatic vehicles before they enter the market (US Department of Transportation, 2016). The policy includes guidance applicable to all intelligent systems on the vehicle, such as privacy, system safety, crashworthiness, consumer education, and training. Further, it also provides guidance specific to different contexts, such as the operation situation, object/event detection and response, and minimal risk condi-

tions. With clear standards and guidelines, governments and businesses can better govern their design process for ethical and fair AI applications.

In addition to compliance with the core principles, the design of AI analytics also needs to follow the specific principles of data privacy, algorithm accountability, and transparency. First, *data privacy* not only implies the rights of the data subjects to know and control how their data could be collected and used, but also the responsibility of businesses and organizations to collect and use data with the consent of the data subjects and protect the privacy of the subjects. For example, the European Union Legal Affairs Committee recommends privacy by design and privacy by default, informed consent, and encryption, as well as use of personal data need to be clarified (Guan, 2018). Moreover, the recently enforced General Data Protection Regulation (GDPR) further puts these principles into law (European Commission, 2016). Data subjects must be clearly informed of the scope of data collection, the legal basis for processing of personal data, the duration of retaining the data, the involvement of third parties, and disclosure of any automated decision-making that is made on a solely algorithmic basis. In addition, data subjects hold the privacy rights to request the information related to their data collection and use.

Second, different from behavior accountability, which refers to the general responsibilities of different parties for outcomes, *algorithmic accountability* indicates the specific responsibility of algorithm designers to provide evidence of potential or realized harms (WWW Foundation, 2017). Particularly, the potential or realized harms could not only refer to legal issues, but also include ethical concerns, such as algorithmic discrimination. Correspondingly, it is the responsibility of the designers to ensure that the algorithm is fair, explainable, auditable, accurate, and responsible. In addition to clarifying the parties accountable for harms or damages caused by algorithmic decision-making, it is also important to explicitly define who are responsible for repairing the systems to avoid future problems. In the U.S., the Algorithmic Accountability Act of 2019 has been recently proposed as the first federal legislative effort to raise awareness of potential negative impacts of implementing AI systems among various industries, such as technology companies, banking, insurance companies, and other consumer businesses (Jones Day, 2019). It authorizes and directs the Federal Trade Commission ("FTC") to issue and enforce regulations that will require certain persons, partnerships, and corporations using, storing, or sharing consumers' personal information to conduct impact assessments and address any identified biases or security issues in a timely manner.

Finally, the AI analytics design also needs to be *transparent*, which requires verification and auditing of datasets and algorithms for fairness, robustness, diversity, nondiscrimination, and privacy (Stoyanovich et al., 2016). Transparency goes beyond the demonstration of the algorithm, and requires that algorithmic decisions as well as any data driving those decisions can be explained to end-users and other stakeholders in nontechnical terms (WWW Foundation, 2017). It is a necessary component to support algorithm accountability, and has been written into several policies and regulations, such as EU's GDPR. GDPR requires that the decisions made by AI applications should be explainable to the data subjects. The challenges here could be to define a clear standard for a satisfying explanation and overcome the inherent difficulties in explaining AI algorithms (Lawton, 2018).

All the above AI policy principles contribute to a safe, fair, accountable, and transparent smart city environment. Such an environment facilitates the deployment of enabling resources (i.e., smart technology, human capital and institutional support) for the actualization of smart city functioning in various domains of city governance and urban life (Gil-Garcia, Zhang, & Puron-Cid, 2016).

Limitations and Future Work

As with other research, this study also entails a few limitations. First, though our framework is not country or region-specific, and could potentially be applied across different countries and cities, the examples we used to illustrate the smart city AI applications, challenges, and principles are mainly from cities in developed countries, such as the EU members, Singapore, and the U.S., which may limit the generalizability of our work. Future research could test the generalizability and adapt the framework to other countries and cities with different levels of economic development, legal systems, cultural norms, and degree of urbanization. For instance, instead of studying existing cities, future work could also explore the principles needed to build new smart cities from scratch in rural or semi-urban areas, which could be important for developing countries, such as India and China.

Second, our framework focuses on the principles that need to be considered by policy makers and companies to develop and implement smart city AI applications, while not comprehensively exploring the mechanisms needed to operationalize these principles and govern such applications. For example, we did not examine how power issues could be addressed between different stakeholders, such as between cities and state or national governments, between government bodies and industry, or even between big tech companies and governments, in order to implement the proposed principles. Future research could identify the governance mechanisms for smart city AI applications and integrate them with our conceptually derived framework to deepen the understanding of smart city AI governance.

Conclusion

In this chapter, we discuss smart city dimensions, AI system components, smart city AI applications and their regulatory challenges, and the principles for regulation of smart city AI. Particularly, AI applications are being implemented to support various smart city functions, while they present multiple legal or ethical issues. However, the current approach to resolve these challenges is fragmented. Therefore, we attempt to identify common principles that could help regulate AI systems for smart cities. This could assist governments and businesses to form a cohesive view toward the design and implementation of AI for smart cities.

Appendix: A Review of Major AI Governance Frameworks

Table 1 Review of major AI governance frameworks (selected from https://algorithmwatch.org/en/project/ai-ethics-guidelines-global-inventory/)

Source/principle	Transparency/ interpretability	Equality/ fairness	Accountability	Safety/ controllability	Fit with human values	Responsible use/ sustainability	Privacy/ data control	Others
(China) National New Generation Artificial Intelligence Governance Expert Committee, Laskai and Webster (2019)	–	Yes	Yes	Yes	• Harmony and friendliness	–	Yes	• Open collaboration • Agile governance
(European Union) European Group on Ethics in Science and New Technologies (2018)	–	Yes	Yes	Yes	• Human dignity and autonomy • Democracy	Yes	Yes	–
(European Union) European Commission (2018)	Yes	Yes	Yes	Yes	• Human agency and oversight • Societal and environmental well-being	–	Yes	–
(Japan) Japanese Society for Artificial Intelligence (2017)	Yes	Yes	Yes	Yes	• Contribution to humanity • Abidance of laws and regulations	Yes	Yes	–
(Singapore) Monetary Authority of Singapore (2018)	Yes	Yes	Yes	–	• Human ethics (e.g. company value, code of conduct)	–	–	–

Table 1 (continued)

Source/principle	Transparency/ interpretability	Equality/ fairness	Accountability	Safety/ controllability	Fit with human values	Responsible use/ sustainability	Privacy/ data control	Others
(Singapore) Singapore Personal Data Protection Commission (2019)	Yes	Yes	–	Yes	• Human centric (well-being)	–	–	–
(UK) UK Government (2019) and The Alan Turing Institute (2019)	Yes	Yes	Yes	Yes	• Human dignity • Connection • Well-being • Social priorities protection	–	Yes	–
Microsoft (2019)	Yes	Yes	Yes	Yes	• Empowerment	–	Yes	–
IBM (2019)	Yes	Yes	Yes	–	• (Company) value alignment	–	Yes	–
PwC (2019)	Yes	Yes	–	Yes	–	Yes	–	• Governance
SAP (2018)	Yes	Yes	–	Yes	• Driven by (company) value • For people	–	Yes	–
Accenture (2016)	Yes	Yes	Yes	Yes	• Human centric	Yes	Yes	–
IEEE (2019)	Yes	–	Yes	–	• Human rights, well-being • Empowerment	–	Yes	• Effectiveness • Awareness of misuse • Competence
International Conference of Data Protection and Privacy Commissioners (2018)	Yes	Yes	Yes	–	• Empowerment	Yes	Yes	–
Information Technology Industry Council (2017)	Yes	–	Yes	Yes	–	Yes	–	• Robust and representative data
Dataethics (2017)	Yes	Yes	Yes	–	• Human centric	Yes	Yes	–

References

Accenture. (2016). Universal principles of data ethics. Retrieved Oct 25, 2019, from https://www.accenture.com/t20160629t012639z__w__/us-en/_acnmedia/pdf-24/accenture-universal-principles-data-ethics.pdf.

Albino, V., Berardi, U., & Dangelico, R. M. (2015). Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of Urban Technology, 22*(1), 3–21.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. Retrieved Oct 25, 2019, from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Balakrishnan, A. (2018). Cloud security concerns in 2018: Data breaches, security misconfigurations, AI and botnets. Retrieved Oct 25, 2019 from https://www.imperva.com/blog/2018-cloud-security-concerns-data-breaches-security-misconfigurations-ai-and-botnets/.

Brisimi, T. S., Cassandras, C. G., Osgood, C., Paschalidis, I. C., & Zhang, Y. (2016). Sensing and classifying roadway obstacles in smart cities: The street bump system. *IEEE Access, 4*, 1301–1312.

Chessen, M. (2017). The AI policy landscape. Retrieved Oct 25, 2019, from https://medium.com/artificial-intelligence-policy-laws-and-ethics/the-ai-landscape-ea8a8b3c3d5d.

Crawford, K. (2016). Artificial intelligence's white guy problem. Retrieved Oct 25, 2019, from https://www.microsoft.com/en-us/research/publication/artificial-intelligences-white-guy-problem/.

Dataethics. (2017). Data ethics principles. Retrieved Oct 25, 2019, from https://dataethics.eu/data-ethics-principles/.

Dirks, S., & Keeling, M. (2009). A vision of smarter cities: How cities can lead the way into a prosperous and sustainable future. Retrieved Oct 25, 2019, from https://www.ibm.com/downloads/cas/2JYLM4ZA.

Etzioni, O. (2017). How to regulate artificial intelligence. Retrieved Oct 25, 2019, from https://www.nytimes.com/2017/09/01/opinion/artificial-intelligence-regulations-rules.html.

European Commission. (2016). Data protection in the EU. Retrieved Oct 25, 2019, from https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en.

European Commission. (2018). Ethics guidelines for trustworthy AI. Retrieved Oct 25, 2019, from https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top.

European Group on Ethics in Science and New Technologies. (2018). Statement on artificial intelligence, robotics and 'autonomous' systems. Retrieved Oct 25, 2019, from http://lefis.unizar.es/wp-content/uploads/EGE_Artificial-Intelligence_Statement_2018.pdf.

Flasiński, M. (2016). *Introduction to artificial intelligence*. New York: Springer.

Gasser, U., & Almeida, V. A. (2017). A layered model for AI governance. *IEEE Internet Computing, 21*(6), 58–62.

Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. Berkeley, CA: University of California Press.

Gil-Garcia, J. R., Zhang, J., & Puron-Cid, G. (2016). Conceptualizing smartness in government: An integrative and multi-dimensional view. *Government Information Quarterly, 33*(3), 524–534.

Girtelschmid, S., Steinbauer, M., Kumar, V., Fensel, A., & Kotsis, G. (2014). On the application of big data in future large-scale intelligent Smart City installations. *International Journal of Pervasive Computing and Communications, 10*(2), 168–182.

GIS Geography. (2018). The rise of machine learning (Ml): How to use artificial intelligence in GIS. Retrieved Oct 25, 2019, from https://gisgeography.com/deep-machine-learning-ml-artificial-intelligence-ai-gis/.

Global Future Council on Human Rights. (2018). How to prevent discriminatory outcomes in machine learning. Retrieved Oct 25, 2019, from http://www3.weforum.org/docs/WEF_40065_White_Paper_How_to_Prevent_Discriminatory_Outcomes_in_Machine_Learning.pdf.

Goodman, B., & Flaxman, S. (2016). European Union regulations on algorithmic decision-making and a 'right to explanation'. arXiv preprint arXiv:1606.08813.

Greene, M. T. (2018). Today in legal artificial intelligence. Retrieved Oct 25, 2019, from https://www.marketintelligencellc.com/tag/singapore/.

Guan, M. (2018). Regulating AI in the era of big tech. Retrieved Oct 25, 2019, from https://thegradient.pub/regulating-ai-in-the-era-of-big-tech/.

Healthcare Asia. (2017). Singapore is turning to artificial intelligence for elderly care. Retrieved Oct 25, 2019, from https://healthcareasiamagazine.com/healthcare/in-focus/singapore-turning-artificial-intelligence-elderly-care.

Heath, N. (2018). What is AI? Everything you need to know about artificial intelligence. Retrieved Oct 25, 2019, from https://www.zdnet.com/article/what-is-ai-everything-you-need-to-know-about-artificial-intelligence/.

Hofman, J. M., Sharma, A., & Watts, D. J. (2017). Prediction and explanation in social systems. *Science, 355*(6324), 486–488.

IBM. (2019). Everyday ethics for artificial intelligence. Retrieved Oct 25, 2019, from https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf.

IEEE. (2019). Ethically aligned design. Retrieved Oct 25, 2019, from https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf?utm_medium=undefined&utm_source=undefined&utm_campaign=undefined&utm_content=undefined&utm_term=undefined.

Information Technology Industry Council. (2017). AI policy principles executive summary. Retrieved Oct 25, 2019, from https://www.itic.org/dotAsset/50ed66d5-404d-40bb-a8ae-9eeeef55aa76.pdf.

International Conference of Data Protection & Privacy Commissioners. (2018). Declaration on ethics and data protection in artificial intelligence. Retrieved Oct 25, 2019, from https://www.privacyconference2018.org/system/files/2018-10/20180922_ICDPPC-40th_AI-Declaration_ADOPTED.pdf.

Japanese Society for Artificial Intelligence. (2017). The Japanese Society for artificial intelligence ethical guidelines. Retrieved Oct 25, 2019, from http://www.ai-elsi.org/wp-content/uploads/2017/05/JSAI-Ethical-Guidelines-1.pdf.

Jones Day. (2019). Proposed algorithmic accountability act targets bias in artificial intelligence. Retrieved Oct 25, 2019, from https://www.jonesday.com/en/insights/2019/06/proposed-algorithmic-accountability-act.

Jucevičius, R., Patašienė, I., & Patašius, M. (2014). Digital dimension of smart city: Critical analysis. *Procedia-Social and Behavioral Sciences, 156*, 146–150.

Juniper Research. (2018). Smart cities – What's in it for citizens? Retrieved Oct 25, 2019, from https://newsroom.intel.com/wp-content/uploads/sites/11/2018/03/smart-cities-whats-in-it-for-citizens.pdf.

Khalik, S. (2015). Hospitals turn to cutting-edge robots and technology for healthcare assistance. Retrieved Oct 25, 2019, from https://www3.ntu.edu.sg/CorpComms2/Documents/2015/07_Jul/ST_150724_Hospitals%20turn%20to%20cutting-edge%20robots%20and%20technology%20for%20healthcare%20assistance.pdf.

Kidal, D. (2017). The robot that could revolutionise home care for elderly people. Retrieved Oct, 2019, from https://lerablog.org/health/elder-care/the-robot-that-could-revolutionise-home-care-for-elderly-people/.

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal, 79*(1), 1–14.

Laskai, L., & Webster, G. (2019). Translation: Chinese expert group offers 'governance principles' for 'responsible AI'. Retrieved Oct 25, 2019, from https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-expert-group-offers-governance-principles-responsible-ai/.

Lawton, G. (2018). GDPR regulations put premium on transparent AI. Retrieved Oct 25, 2019, from https://searchenterpriseai.techtarget.com/feature/GDPR-regulations-put-premium-on-transparent-AI.

Lombardi, P., Giordano, S., Farouh, H., & Yousef, W. (2012). Modelling the smart city performance. *Innovation: The European Journal of Social Science Research, 25*(2), 137–149.

Microsoft. (2019). Microsoft AI principles. Retrieved Oct 25, 2019, from https://www.microsoft.com/en-us/AI/our-approach-to-ai.

MIT Technology Review. (2015). Robotic surgery linked to 144 deaths since 2000. Retrieved Oct 25, 2019, from https://www.technologyreview.com/s/539521/robotic-surgery-linked-to-144-deaths-since-2000/.

Monetary Authority of Singapore. (2018). Principles to promote Fairness, Ethics, Accountability and Transparency (FEAT) in the use of artificial intelligence and data analytics in Singapore's financial sector. Retrieved Oct 25, 2019, from https://www.mas.gov.sg/~/media/MAS/News%20and%20Publications/Monographs%20and%20Information%20Papers/FEAT%20Principles%20Final.pdf.

Nam, T., & Pardo, T. A. (2011). Conceptualizing smart city with dimensions of technology, people, and institutions. In *Proceedings of the 12th annual international digital government research conference: Digital government innovation in challenging times* (pp. 282–291). New York: ACM.

Nonnecke, B., & Newman, J. C. (2019). How the 'california effect' could shape a global approach to ethical AI. Retrieved Oct 25, 2019, from https://www.weforum.org/agenda/2019/10/ai-tech-ethics-california-facial-recognition.

Nyholm, S., & Smids, J. (2016). The ethics of accident-algorithms for self-driving cars: An applied trolley problem? *Ethical Theory and Moral Practice, 19*(5), 1275–1289.

PWC. (2019). A practical guide to responsible artificial intelligence (AI). Retrieved Oct 25, 2019, from https://www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/what-is-responsible-ai/responsible-ai-practical-guide.html.

Roggensack, C. J., & Abrahamson, J. (2016). Loomis V. Wisconsin. Retrieved Oct 25, 2019, from https://harvardlawreview.org/2017/03/state-v-loomis/.

Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: A modern approach*. Malaysia: Pearson Education Limited.

SAP. (2018). SAP's guiding principles for artificial intelligence. Retrieved Oct 25, 2019, from https://www.sap.com/products/leonardo/machine-learning/ai-ethics.html#pdf-asset=940c6047-1c7d-0010-87a3-c30de2ffd8ff&page=2.

Simpson, T. (2016). Rise of the healthcare robots: Five ethical issues to consider. Retrieved Oct 25, 2019, from http://www.cmfblog.org.uk/2016/03/30/rise-of-the-healthcare-robots-five-ethical-issues-to-consider/.

Singapore Personal Data Protection Commission. (2019). A proposed model artificial intelligence governance framework. Retrieved Oct 25, 2019, from https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/A-Proposed-Model-AI-Governance-Framework-January-2019.pdf.

Smith, M. (2016). In Wisconsin, a backlash against using data to foretell defendants' futures. Retrieved Oct 25, 2019, from https://www.nytimes.com/2016/06/23/us/backlash-in-wisconsin-against-using-data-to-foretell-defendants-futures.html.

Stoica, I., Song, D., Popa, R. A., Patterson, D., Mahoney, M. W., Katz, R., …, Gonzalez, J. E. (2017). A Berkeley view of systems challenges for AI. arXiv preprint arXiv:1712.05855.

Stoyanovich, J., Abiteboul, S., & Miklau, G. (2016). Data, responsibly: Fairness, neutrality and transparency in data analysis. In *Proceedings of the 19th international conference on extending database technology*, Bordeaux, France.

Straits Times. (2017). New rules for autonomous vehicles. Retrieved Oct 25, 2019, from http://www.straitstimes.com/singapore/new-rules-for-autonomous-vehicles.

Tech Wire Asia. (2018). Aggregating, automating and augmenting AI for SG's smart city. Retrieved Oct 25, 2019, from https://techwireasia.com/2018/06/aggregating-automating-and-augmenting-ai-for-singapores-smart-city/.

The Alan Turing Institute. (2019). Understanding artificial intelligence ethics and safety. Retrieved Oct 25, 2019, from https://www.turing.ac.uk/sites/default/files/2019-06/understanding_artificial_intelligence_ethics_and_safety.pdf.

Thomson, I. (2015). Robot surgeons kill 144 patients, hurt 1,391, malfunction 8,061 times. Retrieved Oct 25, 2019, from https://www.theregister.co.uk/2015/07/21/robot_surgery_kills_americans/.

UK Government. (2019). A guide to using artificial intelligence in the public sector. Retrieved Oct 25, 2019, from https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector.

Urban Redevelopment Authority. (2017). Creating future cities with self-driving vehicles. Retrieved Oct 25, 2019, from https://www.ura.gov.sg/Corporate/Resources/Ideas-and-Trends/Creating-Future-Cities-with-Self-Driving-Vehicles.

US Department of Transportation. (2016). Federal automated vehicles policy – September 2016. Retrieved Oct 25, 2019, from https://www.transportation.gov/AV/federal-automated-vehicles-policy-september-2016.

WWW Foundation. (2017). Algorithmic accountability. Retrieved Oct 25, 2019, from http://webfoundation.org/docs/2017/07/Algorithms_Report_WF.pdf.