

AI Systems and the Possibility of Moral Decision-making in Smart Cities

Anna Kartasheva
Ph. D. (Philosophy), Research Associate
HSE University
Moscow, Russian Federation
akartasheva@hse.ru

Daria Tomiltseva
Ph. D. (Philosophy), Associate Professor
Ural Federal University
Yekaterinburg, Russian Federation
daria.tomiltseva@urfu.ru

Abstract—Technologies and algorithms combine quantitative data about people and deep insights into the causes of human actions, empowering people to create new artificial agents. The article raises the challenges of implementing normativity in urban spaces equipped with technical systems with artificial intelligence. The improvement of urban life is come through access to technology also the emergence of smart cities is, but not everyone welcomes the predictability and transparency it brings. The introduced artificial agents influence the existing system of normativity in the city, abolishing any norms, transforming, and generating others. Such influence cannot always predict accurately, so it is not possible to expect cultural and religious organizations, individual scientists, or government institutions to finally provide a definite list of moral norms and social rules, which need to implement into technical systems. It is worth noting that if existing norms and regulations are insufficient, then citizens develop their own models of interaction. In addition to rational and effective decision-making, it needs for the technical system needs to earn the trust of the citizens whose daily lives it regulates. The article outlines three caveats for morally correct decision-making mechanisms. First, they should be ongoing measures against distortion. Second, prevention of “social blindness” on the part of developers. Third, controlling the use of techniques that attract users’ attention while being addictive or negatively affecting users in general. These three caveats can be the basis for engineers, social researchers, and urban communities to begin working together conceptually to form morally correct smart city solutions.

Keywords—*smart city, moral city, normativity, moral norms, embodied AI systems*

I. INTRODUCTION

“If we were quite certain that everyone was capable of taking us, and all that we regard as personal to us, at our own valuation ; if, in other words, we were as naive as children, or if, on the other hand, we were all as suggestible and lacking in reserve as some hysterics, we should probably have neither persons nor society” [1, P. 17].

Back in the mid-twentieth century, there was an unbridgeable gulf between quantitative knowledge, i.e., generalizable data about a multitude of people expressed in numbers, and capricious and changeable individual human qualitative characteristics and actions. However, modern artificial intelligence systems, integrated into many aspects of urban life, have fundamentally changed the situation. It is no longer necessary to choose between quantitative

generalizations and qualitative depth: everything is grasped, ordered, calculated, but at the same time — individualized and targeted. With access to technologies and algorithms of appropriate quality, urban life will become predictable and transparent. But for some reason modern city dwellers are not always happy about innovations, and the reason for this lies not in the field of economics or informatics, but... in morality! No, of course, we are not referring to alarmist scares in the spirit of “enslaving machines”, but to problems of a different nature: problems of implementing normativity in urban spaces equipped with technical systems with AI.

II. WHAT IS NORMATIVITY?

Normativity is a set of social prescriptions, moral and value beliefs that set guidelines for human actions and settings of intellectual technical systems, allowing to evaluate reality from the position of “right/wrong, good/bad, acceptable/unacceptable”. Some norms and rules are broadcast formally, like etiquette rules of behavior, while others are read and adopted by people intuitively. However, mere knowledge of existing norms is not enough. It is necessary to understand the relevance: contextual, cultural, economic, historical; it is necessary to compare them with basic ideas of justice and goodness. In today’s smart megacities, this issue becomes extremely important; after all, all of the above-mentioned variety of norms must be rationalized, technicalized, and handed over to the implementation of technical systems with AI, endowing them with the ability to propose, predict, and independently make decisions.

If there were a limited number of norms and they did not change, the challenges facing smart city planners would be relatively simple and solvable. However, and it is important for us to say this, every new technology and every artificial agent widely introduced into the public space has a serious impact on the existing system of normativity in the city, abolishing some norms, transforming and generating others. It should be noted that such influence is not always purposeful, and the trajectories of social adaptation are not always subject to accurate and desirable forecasts. For this reason, it is not possible to expect a hundred percent working list of moral norms and social rules from those communities that are traditionally considered to be “experts”, “guardians” and “researchers”, such as humanities scholars, religious organizations or state institutions engaged in cultural policy.

III. SMART CITY - MORAL STANDARDS

Much has been written about smart cities, even a great deal. Ideally, a smart city is a moral city. Whether we want it or not, the moral core is at the heart of the smart city concept. The well-being of people and the environment is always a matter of ethics, of value choices. In fact, behind every decision made within the framework of the smart city concept we meet objectified morality, manifested in regulations, rules, certificates (etc.) — through this objectification set the norms that make it possible for people to live together and for cities to develop as sustainable socio-technical systems.

However, under conditions of accelerating technological progress, there is a strong tension between the application of innovative technologies and the lack of formed moral norms of their application. Whenever the problem of the moral implementation in the smart cities becomes the discussion issue one could distinguish at least three main players of representation of moral requirements spread around every normative decision:

- the government (local and/or state)
- technology developers
- regular citizens (so as not only individuals, but the local communities of different sort).

If the existing norms and regulations are insufficient and the “smart” technology is already in operation, the situation gets out of control and citizens develop their own models of interaction spontaneously, not always in the way the developers and city authorities would like. Each player has its own aims, benefits, and risks from the processes of smartization (and different in quality resources of influence to provide them). This “disposition of players” one should take into consideration so to escape the one-side reality description and present the insufficient ways of the solving problems. The authors of this article personally tried to match all the three positions, but only in dotted form due to the article limitation.

This situation does not arise out of nothing, but is largely dictated by the existing in urbanism (as a section of economic geography, which deals with the study of problems related to the functioning of urban centers) several ways of considering what a smart city should be [2,3]. On the one hand, this is a technocratic view: the city as a factory can be predicted, which means that it is controllable, predictable and, let us not be afraid to admit it, manipulative — it subordinates to itself all the diversity of human behavior and relations. On the other hand, this is a social constructivist approach, which assumes that all our knowledge and predictions will be incomplete, and any high-tech manipulation will sooner or later encounter unexpected cognitive “gray zones”, forms of social resistance and “reverse manipulation”. In the first case, a smart city is an innovative and technological city (but by prioritizing technology, we lose ourselves, i. e. people) [4, P. 18-19]. The second approach does not reject technology, but, according to this way of considering, a “smart” city is a reasonable, thoughtful (however, focusing on the interests of people, we must clearly understand how to identify and take into account these interests, as well as realize that technology is not all-powerful). The first approach assumes that normativity is subject to technicalization, we just need to find the right algorithms; the second approach advocates moral multiplicity, social pluralism, and consensus-seeking.

Ideally, when designing a smart city, we should listen to the voices of both points of view. The first approach forces technical systems with AI to “overtake” the development of moral norms, the second slows down technical and economic development. But, of course, this methodological gap does not limit the problematization of the relationship between the smart city concept and moral norms — it only sets the general contours of the description of specific problems arising in everyday urban practices.

A. How to reach a consensus

How and where to look for consensus? Social theory has been struggling with this question for centuries. In our opinion, an interesting idea was expressed by B. Barrett, R. Horn and J. Fien, who proposed four dimensions of a moral city [5]:

- Ethical leadership/management,
- Ethical planning,
- Ethical business environment,
- Ethical citizens.

These dimensions represent the urban community at different levels, and suggest building a transparent and honest dialog between them to make joint decisions based on the adoption of a common ethical strategy and values (of course, each dimension of the urban community has its own interests and values). It should be noted that decisions made within the framework of a moral city are “suspended” between deontology, to develop conditions of interactions shared by the majority of citizens and urban communities — that is the “moral face” of the city, and consequentialism to make specific decisions within the established moral framework and to understand (if possible) the conditions of violation of the moral framework. It is conceivable to speculate that AI algorithms could mediate moral consensus by collecting and analyzing many disparate claims and positions and calculating the likelihood of the consequences of decisions. This, in turn, would make the transformation of normative systems more predictable and understandable for urban communities themselves.

B. Moral issues of the smart city

Moral problems of the smart city manifest themselves at three levels: value (what is important? what is right?), process (how are political, economic and social processes organized and regulated?), technological (what technical systems ensure the processes and allow articulating values?). Let us consider the above division on a simple example and try to understand why some actions are done and produced in this way and not in another way. Say, why do janitors in the city of M. remove snow on the playground, but do not remove snow on the crosswalk? We can assume that a certain janitor is lazy, or loves children, but can’t stand pedestrians... all this, of course, is speculation. It is better to pay attention to how the efficiency of this janitor’s work is evaluated. Let’s say a person sends a photo of a cleaned playground in a special application as a report on the work done. He receives money for the completed work in the area of his responsibility, perhaps some points to increase his personal rating. If the crosswalk was not in the scope of the janitor’s duties, then having removed snow where it is not required, he will get nothing, not even a thank you. The application algorithm will not count his successes, the altruistic impulse will not be rewarded.

So, the external form of the city, the extent to which its streets and public spaces are cleared of snow and ice, shows the care of the city authorities about the physical well-being and safety of citizens; the urban infrastructure either allows neglecting the cleanliness of the crosswalk, or on the contrary, does not leave the possibility of inconvenience; finally, the social life of the city reflects the demand for a particular crossing among the citizens (how critical is it that the crossing is not cleared of snow?), the attitude to the janitor (respect, support, or on the contrary, disregard). In this case, the regulation of the processes of cleaning local urban space from snow does not entail social involvement. Meanwhile, an app in which the janitor uploads a report on the work done, in coordination with the utilities, could encourage the "altruistic" behavior of the janitor, which in turn would allow to be more attentive to the interests and well-being of citizens, introducing a new condition of normativity. But for morally correct, and even more, morally beneficial decision-making mechanisms to work in smart cities, we need to consider how and who makes them.

C. How to make morally correct decisions?

Let's imagine that the city is run by a program, a certain "AI manager", never tired, never wasting time, and having a huge stock of information. Will our "AI manager" be able to make morally correct decisions? And if we are faced with the task of developing an "AI manager" to manage a smart city, what moral parameters will we need to consider?

On the one hand, we can "turn a blind eye" to normativity, and focusing only on the technical aspects: the integration of information and communication technologies (ICTs) as well as the Internet of Things (IoT solutions) to manage urban assets and make city management decisions. At the same time, we will face the challenges of controlling, analyzing, planning, data collection, and improving the efficiency of service to the citizens. This is where the slippery slope of morally correct decision-making begins. After all, our "AI manager" must not only make his decisions rationally and "correctly" (whatever that means), but also earn the trust of the citizens whose daily lives he regulates. He will have to collect myriad digital traces, which Sh. Zuboff calls "behavioral redundancy" [6] and make decisions based on the existing cases and rules, and in such a way as to both earn the trust of the citizens and not to become an agent of any of the privileged groups.

- This, in turn, will mean, first, taking measures against bias, since algorithms can reproduce any distortions from the source data, which are difficult to detect if one does not understand the source data and their limitations at the expert level (of course, most end-users of algorithms do not have such knowledge).
- Secondly, prevention of one's own "social blindness", into which the position and social conditions of developers, the "coding elite" according to the expression of J. Burrell and M. Fourcade [7], can lead. Such a position, not only socioeconomic, but also moral and worldview, can have a negative impact on the development of strategies for making fair decisions. Often the developers do not reflect on their own moral foundations: "Despite talking about community and mutual exchange (...) they believed equally fervently in individualism, competition and survival of the fittest" [7].

- Finally, third, in order to increase the level of trust of citizens and make AI tools implemented in the smart city more pleasant and desirable, developers may want to resort to all sorts of tricks (here it is enough to recall the use of addictive methods for developing computer programs, which first appeared in the gambling industry) to attract the attention of users [8, 9, 10]. It is possible that the above three caveats can serve as a basis for the beginning of joint conceptual work of engineers, social researchers and urban communities to form morally correct solutions for a smart city.

IV. CONCLUSION

So, morality is at the heart of the smart city and constitutes the essence of its system of normativity. Moreover, a smart city "doomed" to provide for the good of its citizens, like Sisyphus, will forever strive for the unattainable. The network of normative requirements and social ties will never allow the developers of technical systems with AI to make the right and morally correct decision that will satisfy absolutely all urban communities. But the good news is that new AI developments will make it easier to find social consensus. One way or another.

REFERENCES

- [1] R. E. Park "The urban community as a spacial pattern and a moral order" in *The urban Community. Selected papers from the proceedings of the American Sociological Society. 1925.* Chicago: University of Chicago Press, 1926.
- [2] E. G. Trubina "City in Theory: Experiments of Comprehension of Space". Moscow: NLO, 2010.
- [3] L. N. Petryanina "Urbanistics, economics and ecological condition of urban areas". Penza: PSUAS, 2016.
- [4] "Networks of the city: People. Technologies. Authorities" / Edited by E. Lapina-Kratasyuk, O. Zaporozhets, A. Vozyanov; Moscow: NLO, 2021.
- [5] B. F. D. Barrett, R. E. Horne, J. Fien "The ethical city: a rationale for an urgent new urban agenda" in *Sustainability* 8(11):1197. DOI: 10.3390/su8111197
- [6] Sh. Zuboff "The Age of Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power". New York: PublicAffairs, 2019. URL: <https://avalonlibrary.net/ebooks/Shoshana%20Zuboff%20-%20The%20Age%20of%20Surveillance%20Capitalism%20-%20The%20Fight%20for%20a%20Human%20Future%20at%20the%20New%20Frontier%20of%20Power.pdf>.
- [7] J. Burrell, M. Fourcade "The Society of Algorithms" in *Annual Review of Sociology*, Vol. 47, July 2021. URL: SSRN: <https://ssrn.com/abstract=3903829> or <http://dx.doi.org/10.1146/annurev-soc-090820-020800>.
- [8] T. Hwang "Subprime Attention Crisis: Advertising and the Time Bomb at the Heart of the Internet". New York: Farrar, Straus and Giroux, 2020.
- [9] N. Schüll "Addiction by Design: Machine Gambling in Las Vegas". Princeton, NJ: Princeton Univ. Press, 2012.
- [10] S. Vaidhyanathan "Antisocial Media: How Facebook Disconnects Us and Undermines Democracy". Oxford, UK: Oxford Univ. Press, 2018.