Building Transparency and Robustness of AI/ADM Management in Public Sector

Iwona Karkliniewska[1]

[1]*Foundation Moje Państwo, Nowogrodzka St. 25/37, Warsaw, 00511, Poland*

Abstract

Artificial Intelligence algorithms are widely used by the private and public sectors worldwide. The experience of leading technology countries in the use of AI in public administration and the judiciary, such as the United Kingdom and the United States, points to many advantages and effectiveness of management in decision-making. However, in many cases, the application of algorithms and the implementation of the AI and automated decision making (ADM) management process omit the social and legal aspects of individuals. The European Commission has set itself the ambitious goal of regulating the use of AI/ ADM systems in the digital market by adapting and harmonising the rules in the so-called AI ACT. Work on the harmonisation of the provisions of the regulation has taken two years and is controversial in the context of the protection of human rights and social balance, in particular in public administration. Therefore, the proposed topic requires a comprehensive analysis and discussion in the aspect of an ethical approach of the use of AI/ADM, taking into account the protection of fundamental and civil rights based on the example of the proposed legal and ethical solutions in the European Union.

Key words

artificial intelligence governance, ethics by design, public administration services

1. Introduction

Along with the dynamic development of technology and creation of the digital market, there was a need for legal regulation of the increasingly commonly used artificial intelligence algorithms. This is largely conditioned by the social discussion in the context of the ethical use of artificial intelligence in relation to individuals and society and the impact of technology on well-being and human rights, which arose as a result of a broad political and social debate. The ongoing discussion for the few last years has been conditioned by the specificity of artificial intelligence systems, which have a high potential for development and are often compared to the intelligence of a natural person. For example, the definition of general systems

AIxPA 2022: 1st Workshop on AI for Public Administration , December 2nd, 2022, Udine, IT
EMAIL: i.karkliniewska@yahoo.com (I. Karkliniewska)

of artificial intelligence, in many cases, is difficult to specify both in terms of technology, as well as economic, legal or social aspects. From a technological point of view, generic systems are part of the mostly AI base software, can be also a part of more complex systems, which are widely used in the applications introduced to the market. The use of given solutions is very broad, and the possibility of programming functions and activities is constantly developing based on the implementation of processes of the technology of learning machine systems. The ongoing debate in the European Parliament regarding the proposed solutions to regulate the AI/ADM systems used by the European Commission raises a lot of controversy and doubts among politicians and civic society. The proposed legal regulations do not cover all the possibilities of using artificial intelligence systems. This is a very important aspect in the discussion, which should take into account a comprehensive approach to the use of artificial intelligence systems that will be present at every stage of human functioning and individuals social activity. So far, the proposed legislative solutions covered only regulations concerning the principles of introducing given systems to the market in relation to producers of given systems as well partly users. However, the ongoing work on adjusting the legal solutions for the use of artificial intelligence systems draws attention to the human-centric and social factors. Therefore, artificial intelligence systems should, above all, be reliable, transparent, accurate, secure and in line with the values of the European Union. Additionally, should respect and protect human rights and ethical values. In many cases, individuals, whose personal data are used in the processing or decision-making process are not aware of the processes taking place and do not have reliable and transparent information on the systems concerned. The regulation of these issues is very important in the public sector. Given the predominance of public administrations in terms of decision-making using AI/ADM, it may have a negative effect on a person's welfare and his social, physical and probably emotional position in the future. The use of AI systems by public administration authorities can improve the efficiency, governance and process management as well can be an element of support the human resources work in the organisations. However, it requires a specific approach and analysis already at the programming stage of the implementation of AI/ADM systems in management and decision-making processes, in particular in the areas related to physical person and their personal data processing.

Therefore, it is important to create a comprehensive methodology of using artificial intelligence application and algorithms, which should be managed by natural person at every stage of functioning. Additionally, there are should be possibility allowing to reproduce the algorithm of repeatability of made decisions, which are based on input data and the possibility minimizing the risk and black hole of programming decisions in the future. The application of the systems concerned should be preceded by an impact assessment analysis of users, in terms of the security, transparency and measurability of the systems concerned and their potential influence on the individuals to whom they are applied. Nowadays, we can find many studies, reports and analyses on transparency, reliability, as well as the protection of privacy using artificial intelligence systems prepared by experts and government working groups. Discussions on the ethical aspects of artificial intelligence are also becoming more frequent. The concept of ethics in artificial intelligence can be very relative and dependent due to the approach of stakeholders in the created value chain. When these solutions are used to increase efficiency and profit processes, there will always be an uneven distribution of forces and the predominance of participants in the supply chain of artificial intelligence systems technology solutions. Improving the delivery of technological solutions in the process of functioning and data management, could create the risk of social exclusion, the emergence of bias and the inability to control the development of artificial intelligence systems by physical person. Therefore, public administration authorities should consider in this aspect different approaches of governance and move away from the principles of hierarchical process of management to the ethical and responsible rules for the use of artificial intelligence systems. These should be processed at the stage of programming and developing digital management policies.

2. Results and findings

Decision-making in public administration sector is based mainly on the classical institutional model management of law enforcement, judiciary and administrative decision-making towards citizens [8].

Today, research analyses and discourses identifying used artificial intelligence systems in the public sector as an element of governance process, that helps to build the common good of citizens and as a factor of sustainable social development in the created ecosystem of the digital market. The operationalisation of AI/ADM processes requires the development of a comprehensive management of the mechanism, which will be based on ethical principles, taking into account the impact on the social factor and forming an integral part of the entire AI

system infrastructure in public administration. The management of these systems should be human-centric and inclusive. The management of AI systems should be based on good practices, legal regulation, reliable ethical standards and coordinated cooperation between stakeholders. Therefore, it is important to develop standards and guidelines for the ethical use of AI systems at the stage of programming and creating AI algorithms. A group of experts appointed by the EC has developed guidelines aimed at building ethical designed AI systems and algorithms based on the proposed supply chain, process development and data assessment solutions [1]. First of all, AI systems built on the principle of ethics by design should meet the following criteria respecting human rights and privacy, governance data and personal data protection and be transparent, trustworthy and accountable [1]. At the same time, the control mechanisms of these systems should be possible at every stage of the system life cycle and form coherent logical and functional whole. By creating AI/ADM solutions for public administration services, should be taken into account the principle of sufficient and appropriate use of tool. There for more, the diagnosis of the problem should be solved and provided in a most appropriate solution for specific task of cognitive technology of using data, especially personal and biometric data.

As the solution data creation should contain following elements:
- general ethical standards;
- guidelines on the ethics of created processes in artificial intelligence algorithms and systems;
- methodology of used application of artificial intelligence systems;
- methods and tools used for data development.

Recommendations for the development of guidelines for the reliable and transparent use of AI systems in public administration, in general, should be based on the three main pillars of the functioning of the public sector, namely the implementation of policy and law enforcement [1;4], the organizational environment [9] and taking into account all mechanism of the ongoing process of used models. [7;8]. Well designed AI tool has to be the component of optimal and suitable solution.

Models based on ethical principles should take into account:
- specification of tasks and goals of the given models;
- specification of requirements;
- creating a precise model architecture;
- appropriate processing and collection of data for processing;
- detailed study of the development of the functioning of the model;
- models evaluation and testing.

When artificial intelligence systems are built, possible errors should be also considered at every stage of development to avoid the effect of black box and vendor lock-in.

The use and processing of data should be trustworthy, transparent and in accordance with international and EU legislation in the field of data processing, and comply with all ethical standards. In turn, public administration employees in the process of creating a digital data management strategy and implementing AI/ADM algorithms should properly schedule all activities at the stage of the system purchasing procedure.

3.Good practice

Management process approach based on AI algorithms should be explored and developed at rigorous stages of the introduction of new AI/ADM systems or processes into public administration. Public administration authorities should made the impact assessment of appropriateness of using AI/ADM system. Officials, in regards of capability and the level of knowledge, have to examine the existing data resources and analyze the suitable solution of AI systems introduced to the market or launch development of tailored tool for specific needs and purpose.

By implementing the algorithms in the public administration process, there is a need to involve all the levels of management process to answer the following questions:

1. Is the proposed method of AI/ADM algorithms appropriate for the application?
2. Will it work properly for the purpose used for?
3. Will algorithms used in the proposed context be lawful, safe, transparent and have no risk?
4. Will it be adjustable to the principles of good public administration?
5. Do the used system outputs will be understandable for the officials?
6. Can public body officials explain the mechanism of decision making? Can they change it by facing the presented outputs?
7. Do they trust of the used decision making system?
8. Can it be maintained to the required standards as long as is necessary?

AI based projects introduced by public sector are mainly focus on machine learning systems, which are proceeding big volumes of data, including personal data. Developing friendly and transparent digital ecosystem for society, public administration entities have to introduce entire policymaking governance of AI/ADM systems.

As the result created policy should focus on following areas:

• 'Decision to be adopted' based on ethical frameworks, including transparency, staged public procurement processes to meet the ethical requirements;

• Created decision model and source code guidelines should include including external validation of the quality, compatibility and appropriateness of originating decision-model, source code and/or learning system;

• Guidance for data use, sharing, and storage;

• Guidelines and procedural requirements for decision reviews and systems audits used for decision making outcomes regarding individuals or groups.

Overall approach of AI governance can significantly minimize risk of using machine learning systems.

4. Discussion

Civic organization, the Foundation Moje Państwo developed report by proposing standards needed for credible, transparent and ethical management of AI/ADM algorithms in public administration for further discussion. The aim of the proposal is to emphasize key aspects of efficient AI governance in public sector.

In general, good standards of using artificial intelligence technology should be based on the following principles:

- Assessing the appropriateness of using an AI/ADM system;
- Use of innovation-friendly public procurement modes;
- Introduce dedicated system transparency conditions in agreements with AI/ADM system contractors;
- Introduction of AI/ADM policies or regulations in public institutions;
- Applying explainable AI in the public sector;
- Provide education to officials on how to deal with AI/ADM systems;
- Interdisciplinary collaboration between the public sector and other stakeholders;
- Proactively report on the public sector's use of AI/ADM systems with implications for Citizens [3].

The comprehensive approach in the proposed good standards includes a multi-stage demand analysis of ongoing process, implementation of the policy for the use of artificial intelligence solutions, the system management process, as well as the stage of monitoring of accountability of the applied AI solutions, and their potential impact on a natural person. It also draws attention to the important aspect of building an innovative, friendly and competitive environment, already at the stage of creating public procurement procedure. It points to the advantages of the proposed solution, which allows for trustworthy and reliable use of data, by building expertise and knowledge capability and developing cooperation framework between various stakeholders, including research and academic communities, representatives from business and technology sector. The report can be an universal set of guidelines and good practices example of implementation and efficient management of AI/ADM systems e.g. already launched in Canada, Poland and Great Britain [3].

Presented good standards can be developed and amended, according to the nature of current or future needs of the used AI/ ADM systems in public sector.

5. Acknowledgements

Iwona Karkliniewska is a PhD student at Warsaw University at the Faculty of Management in the Academic Subunit for International Economics in the research area of digital market, technology, innovation and data economy analysis in the CEE and Baltic Sea Region. This research has been funded by the project of Foundation Moje Państwo "Public Policy Discourse on Artificial Intelligence"- supported by the European AI Fund, a joint initiative of the Network of European Foundations (NEF).

6. References

[1] European Commission, DG Research and Innovation. Ethics By Design and Ethics of Use Approaches for Artificial Intelligence, 2021. URL: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-bydesign-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf

[2] L. Floridi, J. Cowls, T. C. King, M. Taddeo, How to design AI for social good: seven essential factors. Sci. Eng. Ethics. 26 (2020) 1771–1796.

[3] Foundation Moje Państwo, Good Standards in the Management Artificial Intelligence Systems and Automatic Decision Making in The Public Sector, 2022. URL: https://mojepanstwo.pl/pliki/ai-adm-management-public-sector.pdf

[4] A. Jobin, M. Ienca, E. Vayeana, The global landscape of AIethics guidelines, Nat. Mach. Intell. 1 (2019) 389-399.

[5] A. Mantelero, AI and Big Data: A blueprint for a human rights, social and ethical impact assessment, Computer Law & Security Review, Volume 34, Issue 4, 2018, pp. 754-772.

[6] J. Morley, L. Floridi, I. Kinsey, A. Elhalal, From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices, Sci. Eng. Ethics, 26 (2020) 2141- 2168.

[7] M. Oswald, J. Grace, Sh. Urwin, C. G. Barnes, Algorithmic risk assessment policing models: lesion from the Durham HART model and 'Experimental' proportionality, Information & Communication Technology Law, Vol 27, No. 2 (2018) 223- 250.

[8] A. Sigfrids, M. Nieminen, J. Leikas , P. Pikkuaho , How should Public Administration Foster the Ethical Development and Use of Artificial Intelligence? A review of Proposals for Developing Governance of AI, Front. Hum. Dyn., 2022.

[9] B. Shneiderman, Bringing the gap between ethics and practice. Guidelines for reliable, safe and trustworthy human – centered AI systems. ACM Trans. Interact. Intell. Syst. 10, 2020, 1 – 31.