Journal of US-China Public Administration, September 2025, Vol. 22, No. 2, 59-70
doi: 10.17265/1548-6591/2025.02.001

Ethical Integration of Artificial Intelligence in Public Administration: Policy and Public Value Creation

Yousif El-Ghalayini
Mohammed Bin Rashid School of Government, Dubai, United Arab Emirates

Artificial intelligence (AI) in public administration presents substantial potential to revolutionize how governments operate, offering significant advancements in areas such as public procurement, service delivery, traffic control, and healthcare management (Dimand et al., 2023). This technological evolution aligns with the growing imperative to improve service quality, make better use of resources, and uphold core principles like equity, transparency, and accountability. Consequently, research on AI in public administration must draw on diverse methodological and analytical approaches. A comprehensive and interdisciplinary perspective is essential not only to assess the effects of these emerging technologies but also to ensure their effective integration into established public service systems that can adapt to today's complex governance needs. Scholars have illustrated how AI can optimize service delivery by automating routine administrative functions (Mehr, 2017; Vogl et al., 2020), personalizing services based on citizen needs (de Sousa et al., 2019; Margetts & Dorobantu, 2019), and enabling predictive planning for future demands (Dekker et al., 2022). This paper highlights how public institutions are increasingly engaging with AI, highlighting the need to harness its capabilities through a public value-oriented lens (Desouza & Fatima, 2023).

Keywords: artificial intelligence, public management, public service values

Introduction

Artificial intelligence (AI) is the ability of developing machines that are capable of performing activities that would otherwise require human intelligence throughout the acts of comprehending, choosing, learning, and learning (Andrews, 2018). They can be defined as intelligent applications in the public administration where automated or semi-automated systems enhance the service delivery, operation efficiency and responsiveness through such means as job automation, data analysis and modelling. Previously, Sousa et al. (2019) asserted that AI came as the advancement of computing machines during the 50s up to big data, machine learning, and even neural networks. In the public sectors, AI identifies areas of healthcare, policing, and policymaking, where issues concerning big data and automated decisions are solved (Busuioc, 2020). As such, AI is quickly becoming a decisive phenomenon in public administration with new prospects for increasing effectiveness, openness, and quality of services (Andrews, 2018; Busuioc, 2020; Sousa et al., 2019).

However, there are several concerns when AI is adopted in public organizations, and this revolves around ethics, fairness, and trust (Wirtz & Müller, 2018; Andrews, 2018; Sousa et al., 2019) that affect productivity and how citizens react to any implementation of AI in any service, mainly due to limited provisions on how they are to be used. As such, the primary goal of this paper is to determine the nature and form of ethical, responsible, and

Yousif El-Ghalayini, Ph.D., associate professor, Mohammed Bin Rashid School of Government, Dubai, United Arab Emirates.

valuable application of AI in public administration with a focus on the opportunity to increase effectiveness of public services with AI technologies to solve emerging challenges. Specifically, some researches like Sousa et al. (2019) present some of the uses of AI like profile matching to suit users into certain characteristics. Hence, the key areas that the paper discusses include the application of artificial intelligent algorithms (set of rules to solve a computational problem) like matching descriptions given by a user to a certain application. This is in relation to algorithmic bias (errors in a computer that lead to wrong outcomes) that can direct the decisions in a certain way due to wrong data input, decision-making traceability, and accountability. In addition, this paper assesses the important factors of policy architectures (processes that guide any policy), managerial leadership, and organizational flexibility in delivering AI systems ethically and efficiently. Hence the main research question is "How can public administration integrate AI technologies to enhance service delivery and operational efficiency, while minimizing algorithmic bias and ensuring transparency, fairness, and accountability in decision-making processes that all users can accept?"

The value of public service in a given jurisdiction affects everyone cumulatively and nobody gets affected by the regulations set in place to govern their use (Svara & Baizhanov, 2018). For example, when one walks into a court to seek justice, everyone else in that region gets treated the same under the same law. Here, the value of public service speaks volume since it is not designed to be biased in whatsoever way. As such, the rules apply in engaging AI tools in complementing and not replacing human decision-making while focusing on responsible AI management and public value delivery (Svara & Baizhanov, 2018; Maragno, Tangi, Gastaldi, & Benedetti, 2022). Moreover, the paper focuses on the advancement in the ethical AI policies that leaders enforce, especially in relation to already established policies.

AI Policy Frameworks

There are many uses of AI in the public sector, including health-care, law and order, and social services, among other uses, and concerns on ethics emerge as a result. Cumming, Saurabh, Rani, and Upadhyay (2024) asserted that it is critical to promote the creation of sustainable ethical AI policy frameworks so that AI technology's utilization is appropriate and constructive in public organizations like the police force (Wirtz & Müller, 2018). Since AI-based systems find applications in making the operations effective and in decision making subjects like health-care, law and order, and social services among others, it becomes the responsibility of the governments to come out with specific solution for the ethical issues that may arise (Sarker, 2022). AI should be policy governed in a transparent and accountable way that does not discriminate as the technology threatens social capital. A good example is having a comprehensive policy framework AI Act that outlines a framework for regulating the use of AI in public administration with special focus on protection of fundamental rights and creation of public values (van Kolfschooten & van Oirschot, 2024).

Having an AI Act is an innovative legislation that puts AI into risk categories depending on their impact on the people and society, and this emerges from public reporting by the people on how AI directly or indirectly has a negative effect on their livelihoods (van Kolfschooten & van Oirschot, 2024; Kusche, 2024). In so doing, this risk-based approach plays a significant role in public administration because applications like facial recognition or predictive policing can effect great changes in citizens' rights; especially when used for predatory marketing or spying by agencies it can lead to breaching of privacy of the individuals or access of data illegally. Therefore, having such an AI Act protects the individuals in the state or region in which the Act applies through defining high risk AI systems and applying the set requirements on them. For example, organizations implementing high

risk AI applications should be under obligation to identify all potential risks, improve the data quality, and where necessary, retain human supervision of the AI results. It minimizes the chances of bias reoccurring in the AI or making decisions that are difficult to explain thus hampering fairness and equality in the provision of public services.

As shown by Alvarez et al. (2024), another problem that comes with the use of AI is fairness, and this can affect certain individuals when used inappropriately. Importantly, fairness is crucial in AI policy frameworks especially as it relates to the possibility of AI systems to worsen existing biased behaviors like racism (Alvarez et al., 2024). As shown in the previous section on the value of public service like courts (Svara & Baizhanov, 2018), those working in these areas use AI to aid them in making some decisions and various examples can show how this can show fairness in play. For instance, where AI is used in the form of predictive policing, it means that the algorithm will assign risk to specific demographic groups based on the past crime record which could be prejudiced. Such instances partially explain why non-discrimination policies must be adopted, and why AI systems should be trained on datasets that are diverse and inclusive (Alvarez et al., 2024; Ferrara, 2023). In the area of use of AI in public organizations, AI Act requires that AI systems be checked for bias such as those directed to marginalized communities and assumptions of their behavior (an example is people of color being associated with crime), while the data used to train the AI systems must be clean, comprehensive, and non-biased. This means that by improving fairness in the development and deployment of these AI systems in the public organizations, it will be possible for the use of technology to work towards solving social issues of prejudice and not deepen them.

Another area of focus when it comes to AI policy frameworks for public organisations is transparency, and governments need to avail the necessary information to the public on how their data are used so as there is transparency. Chan (2023) indicated that the citizens' trust essential to the application of AI systems can be assured if the decision-making process and the underlying reasons for AI-affected actions are comprehensible. With this in mind, Kusche (2024) posited that AI Act mainly stipulates that the outputs of any AI systems must be explainable, and thus public agencies must give reasons for AI technologies' decisions. This requirement is a plus for the accountability and the citizens can review the AI decision made by the public officers. For instance, in those welfare distribution systems where the use of AI in decision-making is adopted, there is the need for decisions made to be transparent to avoid making wrong decisions, or making decisions that are prejudicial to the citizens. Without this transparency, the complexity of algorithms that underlie AI exposes the risk of diminishing the society's trust in government services.

Other values of ethical AI policy include fairness and transparency besides accountability which plays a significant role in regulation of artificial intelligence systems (Chan, 2023; Kusche, 2024). A composite AI Act also requires public organisations to retain human supervision in the use of the AI systems, including control of the ultimate decision-making process by human personnel. To this end, this requirement helps prevent the AI systems from making decisions that would have a profound impact on the people's lives. For instance, the automatic decision-making systems involving the ways that people obtain housing, welfare benefits, and other social services can only be luminously monitored by humans so that people do not suffer from the AI's prejudice decisions. Such measures through the AI Act mean that AI systems are to operate solely as the tools that combine and strengthen human control.

The issue of appropriate ethical AI policy frameworks deserves attention in order to direct appropriate AI usage in the public sector. It is vital to focus on equal treatment, non-discriminatory, and unbiased approach, by

making the principles of the fairness fully transparent and accountable. A positive scheme is through AI Act that provides a strong regime that sets out the future guidelines on the ways AI technologies have to be deployed with regard to standards and values (Saheb & Saheb, 2024). As even more public organisations implement such AI systems, it is important for such frameworks to develop to respond to new problems like bias and the lack of transparency in operations thus ensuring that AI works for the benefit of the public whilst also maintaining ethical principles.

AI—Algorithmic Bias

Algorithmic bias is emerging as one of the biggest barriers in implementing AI in public service. It is seen that AI systems work on algorithms that make decision based on the data fed into them and thus when the data used to develop such algorithms are incongruent, then the decisions generated by the AI algorithms will further the inequalities between already existing prejudice regarding certain groups of people in the society like the association of crime with people of color and the judgments made by courts, which is a key component of public service (Akter et al., 2021). This problem is quite relevant in the field of public administration where decision-making is often implemented through an AI algorithm in areas like social services, law enforcement, or healthcare for instance, and algorithmic bias may result in giving disadvantaged groups an unjust treatment. It is equally important to understand how these biases are formed and to develop the ways of preventing this as the AI systems used in public administration ought to bring justice rather than strengthen the prejudice. Therefore, the main concern should be with the data we use to train AI systems, and AI systems designers need to be as fair as possible.

An example of algorithmic bias which is demonstrated in public administration is in the criminal justice system where predictive policing algorithms have been employed to find areas with high propensity to crime (Bignami, 2022). Often such algorithms are based on the crime history, which can contain biases in policing for years. For example, if some ethnic groups have been targeted by police because of racial bias, the algorithm will increase their link with crime when in fact the crimes are the same as other regions. Therefore, it becomes a self-fulfilling prophecy and the predictive policing tools might arrest mainly the minorities. Hence, the data used need to be critically evaluated to eliminate such bias. But in order to find a solution to this issue, the public organizations evaluate the data that were used in training of the AI systems and are certain that it is not influenced by the old biases.

Further, the same issue can be identified when it comes to AI application in social services where it has been seen that AI systems have been applied in the context of calculating the welfare benefits or housing subsidies (Reamer, 2023). Nonetheless, if the learned data contain prejudices of socio-economic status, race, or any other attribute, then the system will serve certain groups of people even more unfair. For instance, an AI system can bring out the correlation between low-income earners with fraud and hence apply higher scrutiny to their applications than those of higher income individuals. It increases current disparities of opportunities and it erodes the people's confidence in government services. To avoid the above biases, there are measures that public organizations must embrace such as auditing of the algorithms to check for bias and the data feeding the algorithms must be bias free.

Addressing bias in algorithms is not a one step process. Reamer (2023) showed that different strategies need to be implemented to address the challenge of algorithmic bias. One of them is to garner that AI systems are trained on datasets that are inclusive of all the different categories of people. Public organisations, to be specific, must make sure the data they gather are rich and diverse and do not over-rely on data profiles which are dated

and which are usually created with already existing biased perceptions (Hodgson, Watts, & Gair, 2023). For example, if an AI system is being created to determine the extent of people eligible for housing benefits, the data required should be correctly reflecting people of different income status, race, location among others to avoid bias. Further, there are other strategies like re-shuffling or re-sampling which an organization can employ to try and balance the training data.

There is also the necessity for constant auditing of AI systems as some of the identified strategies. These audits, as Minkkinen, Laine, and Mäntymäki (2022) suggested, can evaluate the effectiveness of the algorithm within the multiplicity of demographic categories with a view of realigning decision-making. For example, when an AI tool in the police station yields different outcomes in various racial or ethnic groups, it will suggest that the system is perhaps biased. For instance, it is only through regular auditing that biases of the public organizations can be identified and corrected at an early stage. However, there is a lack of transparence in AI systems; public organization decision should be more transparent and explain why they are making such decision and the affected individual should be able to challenge the decision if they find it unjust.

There is also a need for public organisations to encourage ethical practice and development of AI both for developers and administrators by providing them with knowledge on how bias affects decision-making (Stahl et al., 2021). This includes obliging developers to incorporate the ethical aspect of their systems during the development stage and also promoting the use of bias elimination methods like fairness constraints for machine learning. For example, developers who want to implement an algorithm for hiring can employ something like fairness-aware algorithms that change the decision-making pattern they are to pursue in order to make it impossible for some categories of people, such as minorities or veterans, to be discriminated against.

Therefore, the problem of algorithmic bias remains a major issue affecting the proper utilization of AI in the public administration (Stahl et al., 2021; Minkkinen et al., 2022). Bias in data can channel to unfair consequences to the population especially those who are marginalized in the society and are counterproductive to the advanced use of AI for better provision of public services. These concerns mean that public organisations require the following measures to address these concerns, utilizing diverse datasets, conducting frequent audits, enhancing transparency, and promoting ethically aligned AI. If these issues are going to be solved, the public administrations can be benefit by using AI solutions while being fair and respecting the public.

Public Service Values

Transparency and accountability are some of the principles of public administration and governance particularly when it comes to using AI for decision making. The problem is that as AI-based systems find their way into the provision of government services, the reasons as to why certain decisions have been made are also important and that is why the decision-making of such systems has to be modelled carefully (Cheong, 2024). When used in welfare distribution, policing, or medicine, an algorithm can make life-changing decisions on citizens they are not even aware of. Failure to make these systems transparent may result in their being turned into effectively closed systems, or “black boxes”, in which the decision-making process itself becomes difficult to determine or understand, or where decisions themselves can be manipulative or unfair. To avert this, the AI development processes must incorporate points where the human performance can intervene and be checked by the public.

An example of a situation where transparency is required is in employment of AI systems to assess qualification to receive welfare benefits. The government, for instance, has in the recent past used AI tools in

determining applicants for social benefits like unemployment benefits or housing subsidies (Gesk & Leyer, 2022). However, whenever decisions are made purely in this statistic mathematical form rather than given a brief explanation to the applicant, then it feels like discrimination. One of the real-life incidents is in the Netherlands where due to the implementation of AI-based system for welfare fraud, benefit claims were selectively denied from specific ethnic groups (Maxwell & Vamparys, 2020). In another case, the absence of information disclosure on how the AI arrived at the decisions provided caused a lot of public outcries and the program was closed. This example supports the idea of making exposures of AI systems not only in terms of their working mechanism but also, the output of such systems needs to be explainable to citizens so that they can scrutinize the same.

Thus, problems describing explainability of AI systems should be solved in order to make the decisions made by AI systems transparent in public organizations (Maxwell & Vamparys, 2020). This is in light of the fact that the rationale behind AI's decision-making should be easily understandable. For instance, if an AI system is being applied in evaluating loan requests or the applicants for public housing, then it must be in a position to explain why applicant "X" has been granted credit while applicant "Y" was rejected the same. Explanations generate accountability since it enables those who made the decisions as well as the other victims to understand the thinking of the AI (Nallakaruppan et al., 2024). In addition, transparency is the protection against error and preconception; where the information technology is open to the public, there are more chances of recognizing failures or unfairness in the program.

Transparency is also well related to accountability, as mentioned above. Public organizations, thus, have to assure that none of the AI are completely autonomous decision-makers but are at least supervised by humans (Gesk & Leyer, 2022). For instance, in the criminal justice system, there is such an AI tool as a risk assessment which includes algorithms that help in decisions concerning the granting of parole or the pronouncement of the sentence. These tools compute the risk of recidivism depending on the provided factors; however, the effectiveness of these tools has been criticized, especially with scenarios, where the generated racial bias is detected by the algorithms. Whenever this is done, there is a need to have human intervention in order to assess the algorithm solution and then make an ethical decision. To avoid AI explanations for sale being given as a scapegoat of decisions, then accountability means that public decision-makers remain accountable for such decisions made with the assistance of AI explanations for sale.

Another way through which public trust in AI-generated government services can be improved is through independent and periodic check on AI systems to make sure the system is periodically reviewed to address any bias or fairness issues if they are present so that the services rendered are equal and fair to everyone (Akter et al., 2021). For example, if a machine learning-based AI application for public health was promoting denial of healthcare benefits to a particular race, then an audit can make the organization come across this vice and work toward fixing it. Also, sharing the outcomes of such audits encourages trust because people can notice the government's efforts toward ethical artificial intelligence. Furthermore, there are legal initiatives in AI Act that aim at enhancing transparency by requiring companies to disclose numerous documents regarding the functionality, decision making, and risk evaluation of the high-risk AI systems that are used in public administration (Kusche, 2024). Such regulations are vital especially for the aspects of accountability since it makes it impossible for public organisations to implement the AI systems without having to go through the various issues of ethics without being in a position to explain to the public given sufficient publicity.

AI Paradigm Shift

It is important for the public sector, mainly controlled by the government to consider and to take much care while implementing the use of AI in the public sector (Bignami, 2022). In the current and future course of development of AI technologies as vehicles for public administration, sound leadership is imperative for the right utilization of the technologies in the public domain within the context of ethical applications, public values, and trust in government. Decision-makers in the public sector have responsibilities in the implementation of AI systems in a way that shall not compromise or remove executive control over the decision-making process.

It is the duty and social responsibility of the public leaders to guarantee the use of AI as support tools to augment the human discretion and not tools to displace human discretion. For instance, in the areas such as healthcare, such systems help doctors in the identification of diseases by studying numerous pieces of information and tracking several patterns which might be unnoticeable to humans. Thus, it is critical to state that AI should not be the decision-maker when it comes to creating the treatment pathways of patients. However, leaders in the public sector must create paradigms which entail that AI is helpful in the sense that its present findings are helpful in the procedures; but the doctors are the ones who make decisions (Ferrara, 2023). This is important so that there is a balance that can help in sustaining accountability and also guarantee that human experience remains important when making important decisions.

Managers in public organisations like courts and public healthcare also cannot avoid the ethical issues related to the implementation of AI decision-making which include issues of bias, fairness, and transparency (Reamer, 2023; de Manuel et al., 2023). Managers are supposed to act as to set the rules for AI disclosure within their facilities and make sure that the AI devices will not have bias and discrimination. For example, in the policemen service, the AI based facial recognition technologies have been adopted to increase rate of observation and suspect identification. Nevertheless, de Manuel et al. (2023) posited that these systems tend to be racist against certain groups and may lead to wrong identification. The directors of the public sector have to make sure that these AI tools are examined and audited so that such biases are not present within, and used properly (Minkkinen et al., 2022). It is clear that through risk management of ethical risks the leaders can enable the adoption of AI within their organisations in a manner that is conducive to trust from the public.

Also, leaders have to ensure that cooperation between an AI system and employees occurs by stating that AI is a tool created to enhance human resources (Arslan et al., 2021). In the public sector for instance where the impacts resulting from decisions made are often ending up being more significant, such an approach is crucial. For instance, AI systems can be applied in public welfare programs to determine the participants' eligibility for certain benefits based on their financial records and other factors. However, the public sector managers need to make sure that human beings are the last word in deciding the patients' eligibility especially in situations where compassionate stories need to be told and heard. In this way, the leaders will be able to design a synergy between AI and human workers to signify how AI can augment the judgment and improve the decision-making with the advantage of fairness.

Of equal significance is the concern on the aspect of education and skills development as the leadership takes charge of the AI adoption processes (Fullan, Azorín, Harris, & Jones, 2023). Lack of skill and competency in the human resource in the public sector is a factor that limits the proper utilization and management of AI technologies; thus, it is imperative for public sectors to spend on the training of the human resource to properly utilize AI technologies. This includes not only the kind of skill needed to "speak" to machines but also the kind

of experience needed in order to grasp the constraints and the possible ethically questionable aspects of the AI. The leaders who employ AI in decision-making processes concerning public services must take an extra measure to ensure that the staff within the organizational possess the necessary skills in the interpretation of AI generated outputs and the human-centered judgement in unexpected circumstances (Alasmri & Basahel, 2022). For instance, in the public education systems, these innovative instruments can help the teacher recognize the extra attention required students based on the performance indicators. Nevertheless, several decisions regarding schooling still remain exclusive to the teacher's discretion based on their professional content knowledge and professed-general pedagogical experience of students.

Public Servants and AI

Since the use of AI technologies becomes widespread in the framework of public administration and delivery of public services, requalification and development of competencies of the civil servants become highly relevant and important (Yang, 2022). Thus, they would be able to effectively manage and govern AI systems and technologies. AI will bring about enormous returns in terms of productivity, optimal decision-making, and quality service delivery provision. Nevertheless, to realize the expected gains and benefits, AI requires a competent workforce that appreciates both the technology interface and the sundry ethical, legal, and operational considerations (Jaiswal, Arun, & Varma, 2021). Training and skills development for employees in the public sector is critical for responding to AI solutions, not forgetting improving on career opportunities for workers with an aim of avoiding displacement of their positions by the technologies.

Algorithm-based decisions, repetitive entails like clerical, data entry or document processing, administrative processing, or decision-making can be performed through AI application which makes people afraid of loss of job (Yang, 2022; Na, 2024). Although it has been seen that AI can replace human employee, there is an optimistic aspect that AI can perform the repetitive task and free the human employee for the value-added tasks. For instance, AI in public health administration leads to data dealing and patient files processing and data-driven duties and responsibilities and allows the professionals to serve the patients and to engage in critical decision making. Retraining of the labour workforce is mandatory because they all have to shift from performing mechanical tasks to those that call for critical thinking, solving problems, and making decisions that human beings alone are not capable of making.

For this to happen, the public sector organizations have to embark on extensive retraining sessions for their staff to develop skills that will enable them to effectively interact with AI systems (Na, 2024). Technical training is one with the help of which employees would know how to use AI, how to read AI results, and how to solve AI-related technical problems. For instance, AI application is being widely used in the prediction of crimes and facial recognition in law enforcement. Police officers as analysts, also need to be educated on not only the use of the systems but also their drawbacks and the bias influences on the AI predictions (Saheb & Saheb, 2024). This clearly is a technical requirement but the need to decide whether to use AI for such crucial issues as security and protection of the public entails incorporating an ethics factor.

The new competencies which need to be advanced for public sector employees include data literacy and AI governance. For AI to work, there should be proper ways through which data are collected and managed and proper analysis done by employees in the organization (Janssen et al., 2020). This pertains to awareness of legislators' provisions, security standards, and ethical issues with regard to the utilization of personal data. For example, in social services, workers who employed AI systems for evaluating the respondent's eligibility for

benefits must guarantee the quality of data that the AI systems use. Furthermore, employees should also know ways by which citizens' privacy can be defended mostly when involving sensitive data (Na, 2024). Retraining programs should therefore comprise of data ethics and the legal requirements of the use of AI as well as data in the public administration.

Another important aspect of the retraining of the workforce is to follow such concepts as AI integration with human expertise. The public sector employees need to understand how to read AI output in the light of their other professional experience and knowledge (Simón, Revilla, & Jesús Sáenz, 2024). For example, in public education, there exist techniques that may help analyze the student performance data and pinpoint learners who may be performing poorly in their academic work. There is, however, reasonable expectation that in a given context, the teacher must make a professional decision about the feasibility of the recommendation based on their understanding of the learner's needs and realities. This will entail the need to convince the employees to change their mind set about AI as a system that does not require human input all the time. Therefore, as a retraining and competency improvement measures for the workforce involved in the utilization and management of AI in public administration, there is the need for increased investment (Johnson, Albizri, Harfouche, & Fosso-Wamba, 2022). Through employing technical trainings and data literacy, ethical awareness, and collaboration between AI as well as human judgment, public organizations are assured that their personnel is capable of applying AI technologies in a manner that will elongate their useful lives and serve the public's best interest. The proactive approach to reskilling would in the long term help AI become a means through which the public sector becomes empowered, not threatened by its existence.

Substitutability of technologies may also be necessary to bring AI systems in public institutions hence organisational flexibility. With the integration of AI to decision-making and service delivery models, public organizations require changes in several work models, process, and structures (Zirar, Ali, & Islam, 2023). This entails strategic intervention which would involve redesigning organizational structures, operationalizing new work processes to harness potential of AI, building AI culture of innovation, and embedding the right checks and balances that would ensure that AI is used ethically in organizations (Zirar et al., 2023). However, the problem is not only by the introduction of new technologies, but in changing organizational practices to increase the overall efficiency, while not compromising the organization's accountability to the public.

AI systems can also pose challenges for public organizations because the organisations need to change their structure to incorporate the new systems. It usually entails the establishment of the AI governance, management, and execution which sometimes means formation of new dedicated positions and offices (Zirar et al., 2023). For instance, a number of governments are forming AI ethics governance committees or data science committees for overseeing and coordinating the AI projects, for overseeing the performance of the algorithms, and for ensuring that the usage of AI is not unauthorized or improper. Such specialized units aim at making sure that AI systems uphold democratic values such as unfairness, transparency, and accountability (Alvarez et al., 2024). These units have been created based on the realization that AI expertise must be integrated across large organisations such that the adoption of AI in public organisations is properly coordinated within an AI strategy that is informed by expert advice.

Innovation and Regulation in the Public Service

The advancement and adoption of AI have brought in many opportunities for development especially on the way processes are made, decisions are arrived, and services are delivered. But there are also important and subtle

ethical and societal implications associated with the adoption of AI in government services (Kulal et al., 2024). This becomes a challenge where public organizations must find a middle ground in advocating for AI development as well as adopting requisite regulations on implementation of the same. It is aimed at promoting further development of technology while preserving people's confidence and ensuring honesty and righteous behavior.

The intensity of progress in this technology also breeds various shortcomings which call for control through legislation. Enabling AI to take decisions on its own comes with a range of problems, including bias in the decisions made, violation of data privacy, and lack of transparency in the whole decision-making process (de Manuel et al., 2023). For example, in the United States predictive policing system has been accused of promoting racism; it targets minority group because the statistics on which it bases its choices are racially biased (Mugari & Obioha, 2021). Like the welfare automation systems of Australia (Hodgson et al., 2023), there are others that have also been criticized for rehearsing prejudice against citizens that are already marginalized by owing to errors made by algorithms involved in the operations of the systems. These examples illustrate the risks of unregulated advance in AI that describes the potential benefits of utilizing AI technologies at the cost of fairness, justice, and confidence in public organizations. Specifically, this is through the integration of new innovations in the AI field and proper use of the AI for fair and equal treatment of the people.

To solve these ethical challenges, it is important to have regulatory standards to provide guidance when using AI (Minkkinen et al., 2022). Over the past few years, governments and organizations of the international level have approved legislation that can regulate the risks associated with AI, and at the same time promote the use of new technologies. One example is where AI Act proposes a risk-based approach to the regulation, which divides AI systems by risk to human and society (Kusche, 2024). AI that is deemed to be high risk including use in policing or in the healthcare sector will need transparent, accountable, and supervised by man. This favorable oriented regulation makes sure that the AI systems needed for the crucial sectors are strictly under the ethical measurements. However, the executive systems for the other immature sectors are inflated by innovation. The AI Act can be used as an example of how it is possible to liberalize the AI market while still regulating it for people's benefit.

Therefore, the most suitable and mechanisms will suitable and efficient framework of regulation is the one that is flexible and adaptive (Sarker, 2022). Instead, governments need to develop regulatory standards capable to respond to ethical risks, though being open for development along with appearing IT breakthroughs. Examples include instances such as sandboxes where AI technologies can be practiced in enclosed settings but with the convenience of practicing the innovation. Some other key stakeholders are the regulatory bodies that may maintain continuous engagements with AI developers and public organizations to ensure that recommendations for the same from this paper are developed so as to address the issues of responsibility without inadequate restrictions on AI innovation.

References

Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D'Ambra, J., & Shen, K. N. (2021). Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management, 60*, 102387. Retrieved from https://doi.org/10.1016/j.ijinfomgt.2021.102387

Alasmri, N., & Basahel, S. (2022). Linking artificial intelligence use to improved decision-making, individual and organizational outcomes. *International Business Research, 15*(10), 1-15. Retrieved from https://doi.org/10.5539/ibr.v15n10p1

Alvarez, J. M., Colmenarejo, A. B., Elobaid, A., Fabbrizzi, S., Fahimi, M., Ferrara, A., … Ruggieri, S. (2024). Policy advice and best practices on bias and fairness in AI. *Ethics and Information Technology, 26*(2), Article 31. Retrieved from https://doi.org/10.1007/s10676-024-09746-w

Andrews, L. (2018). Public administration, public leadership and the construction of public value in the age of the algorithm and "big data". *Public Administration, 97*(2), 296-310. Retrieved from https://doi.org/10.1111/padm.12534

Arslan, A., Cooper, C., Khan, Z., Golgeci, I., & Ali, I. (2021). Artificial Intelligence and human workers interaction at team level: A conceptual assessment of the challenges and potential HRM strategies. *International Journal of Manpower, 43*(1), 75-88. Retrieved from https://doi.org/10.1108/ijm-01-2021-0052

Bignami, F. (2022). Artificial intelligence accountability of public administration. *The American Journal of Comparative Law, 70*(Supplement_1), i312-i346. Retrieved from https://doi.org/10.1093/ajcl/avac012

Busuioc, M. (2020). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review, 81*(5), 825-836. Retrieved from https://doi.org/10.1111/puar.13293

Chan, C. K. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education, 20*(1), Article 38. Retrieved from https://doi.org/10.1186/s41239-023-00408-3

Cheong, B. C. (2024). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics, 6*, 1421273. Retrieved from https://doi.org/10.3389/fhumd.2024.1421273

Cumming, D., Saurabh, K., Rani, N., & Upadhyay, P. (2024). Towards AI ethics-led sustainability frameworks and toolkits: Review and research agenda. *Journal of Sustainable Finance and Accounting, 1*, 100003. Retrieved from https://doi.org/10.1016/j.josfa.2024.100003

De Manuel, A., Delgado, J., Parra Jounou, I., Ausín, T., Casacuberta, D., Cruz, M., … Puyol, A. (2023). Ethical assessments and mitigation strategies for biases in AI-systems used during the COVID-19 pandemic. *Big Data & Society, 10*(1), 1-11. Retrieved from https://doi.org/10.1177/20539517231179199

Ferrara, E. (2023). Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci, 6*(1), 1-15. Retrieved from https://doi.org/10.3390/sci6010003

Fullan, M., Azorín, C., Harris, A., & Jones, M. (2023). Artificial intelligence and school leadership: Challenges, opportunities and implications. *School Leadership & Management, 44*(4), 339-346. Retrieved from https://doi.org/10.1080/13632434.2023.2246856

Gesk, T. S., & Leyer, M. (2022). Artificial intelligence in public services: When and why citizens accept its usage. *Government Information Quarterly, 39*(3), 101704. Retrieved from https://doi.org/10.1016/j.giq.2022.101704

Hodgson, D., Watts, L., & Gair, S. (2023). Artificial intelligence and implications for the *Australian Social Work* journal. *Australian Social Work, 76*(4), 425-427. Retrieved from https://doi.org/10.1080/0312407x.2023.2247833

Jaiswal, A., Arun, C. J., & Varma, A. (2021). Rebooting employees: Upskilling for artificial intelligence in multinational corporations. *The International Journal of Human Resource Management, 33*(6), 1179-1208. Retrieved from https://doi.org/10.1080/09585192.2021.1891114

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy artificial intelligence. *Government Information Quarterly, 37*(3), 101493. Retrieved from https://doi.org/10.1016/j.giq.2020.101493

Johnson, M., Albizri, A., Harfouche, A., & Fosso-Wamba, S. (2022). Integrating human knowledge into artificial intelligence for complex and ill-structured problems: Informed artificial intelligence. *International Journal of Information Management, 64*, 102479. Retrieved from https://doi.org/10.1016/j.ijinfomgt.2022.102479

Kulal, A., Rahiman, H. U., Suvarna, H., Abhishek, N., & Dinesh, S. (2024). Enhancing public service delivery efficiency: Exploring the impact of AI. *Journal of Open Innovation: Technology, Market, and Complexity, 10*(3), 100329. Retrieved from https://doi.org/10.1016/j.joitmc.2024.100329

Kusche, I. (2024). Possible harms of artificial intelligence and the EU AI act: Fundamental rights and risk. *Journal of Risk Research*, 1-14. Retrieved from https://doi.org/10.1080/13669877.2024.2350720

Maragno, G., Tangi, L., Gastaldi, L., & Benedetti, M. (2022). AI as an organizational agent to nurture: Effectively introducing Chatbots in public entities. *Public Management Review, 25*(11), 2135-2165. Retrieved from https://doi.org/10.1080/14719037.2022.2063935

Maxwell, W., & Vamparys, X. (2020, April 16). *Netherlands welfare case sheds light on explainable AI for AML-CFT (Compliance & Enforcement)*. New York University of Law. Retrieved from https://wp.nyu.edu/compliance_enforcement/2020/04/07/netherlands-welfare-case-sheds-light-on-explainable-ai-for-aml-cft/

Minkkinen, M., Laine, J., & Mäntymäki, M. (2022). Continuous auditing of artificial intelligence: A conceptualization and assessment of tools and Frameworks. *Digital Society, 1*(3), Article 21. Retrieved from https://doi.org/10.1007/s44206-022-00022-2

Mugari, I., & Obioha, E. E. (2021). Predictive policing and crime control in the United States of America and Europe: Trends in a decade of research and the future of predictive policing. *Social Sciences, 10*(6), 234. Retrieved from https://doi.org/10.3390/socsci10060234

Na, S. R. (2024). Application of artificial intelligence in employee training and development. *Mathematical Modeling and Algorithm Application, 1*(1), 26-28. Retrieved from https://doi.org/10.54097/gg5eemnb

Nallakaruppan, M. K., Balusamy, B., Shri, M. L., Malathi, V., & Bhattacharyya, S. (2024). An explainable AI framework for credit evaluation and analysis. *Applied Soft Computing, 153*, 111307. Retrieved from https://doi.org/10.1016/j.asoc.2024.111307

Reamer, F. G. (2023). Artificial intelligence in social work: Emerging ethical issues. *International Journal of Social Work Values and Ethics, 20*(2), 52-71. Retrieved from https://doi.org/10.55521/10-020-205

Saheb, T., & Saheb, T. (2024). Mapping ethical artificial intelligence policy landscape: A mixed method analysis. *Science and Engineering Ethics, 30*(2), Article 9. Retrieved from https://doi.org/10.1007/s11948-024-00472-6

Sarker, I. H. (2022). AI-based modeling: Techniques, applications and research issues towards automation, intelligent and Smart Systems. *SN Computer Science, 3*(2), Article 158. Retrieved from https://doi.org/10.1007/s42979-022-01043-x

Simón, C., Revilla, E., & Jesús Sáenz, M. (2024). Integrating AI in organizations for value creation through human-AI teaming: A dynamic-capabilities approach. *Journal of Business Research, 182*, 114783. Retrieved from https://doi.org/10.1016/j.jbusres.2024.114783

Sousa, W. G., Melo, E. R., Bermejo, P. H., Farias, R. A., & Gomes, A. O. (2019). How and where is artificial intelligence in the public sector going? A literature review and research agenda. *Government Information Quarterly, 36*(4), 101392. Retrieved from https://doi.org/10.1016/j.giq.2019.07.004

Stahl, B. C., Antoniou, J., Ryan, M., Macnish, K., & Jiya, T. (2021). Organisational responses to the ethical issues of artificial intelligence. *AI & Society, 37*(1), 23-37. Retrieved from https://doi.org/10.1007/s00146-021-01148-6

Svara, H. J., & Baizhanov, S. S. (2018). Public service values in NASPAA programs: Identification, integration, and activation. *Journal of Public Affairs Education.* Retrieved from https://doi.org/10.1080/15236803.2018.1454761

Van Kolfschooten, H., & van Oirschot, J. (2024). The EU artificial intelligence act (2024): Implications for healthcare. *Health Policy, 149*, 105152. Retrieved from https://doi.org/10.1016/j.healthpol.2024.105152

Wirtz, B. W., & Müller, W. M. (2018). An integrated artificial intelligence framework for public management. *Public Management Review, 21*(7), 1076-1100. Retrieved from https://doi.org/10.1080/14719037.2018.1549268

Yang, C.-H. (2022). How artificial intelligence technology affects productivity and employment: Firm-level evidence from Taiwan. *Research Policy, 51*(6), 104536. Retrieved from https://doi.org/10.1016/j.respol.2022.104536

Zirar, A., Ali, S. I., & Islam, N. (2023). Worker and workplace artificial intelligence (AI) coexistence: Emerging themes and research agenda. *Technovation, 124*, 102747. Retrieved from https://doi.org/10.1016/j.technovation.2023.10274