



Ethical and Governance Frameworks for Ensuring Safe and Responsible Artificial Intelligence (AI) Development and Deployment in the Public Service

Richard Douglas Kamara ^{a,1,*}
^a University of Johannesburg
richard.douglas_dr@douqualiakademie.org*
* Richard Douglas Kamara

(Received March 1, 2024 Revised April 21, 2024 Accepted May 1, 2025, Available online May 1, 2025)

Abstract

Globally, public entities worldwide face significant service delivery challenges, particularly in underserved communities. These challenges include inadequate access to essential services such as water, electricity, sanitation, housing, and infrastructure. Governments worldwide are actively working towards enhancing service delivery and improving the quality of life for their citizens, and they are increasingly recognising the transformative potential of artificial intelligence (AI) technologies. Nevertheless, implementing artificial intelligence in public service brings forth significant ethical considerations and potential hazards. Basic understanding of ethical strategies and guidelines for implementing artificial intelligence in public service is still in its early stages of development. This paper delves into the potential risks and ethical considerations associated with implementing AI in public service, with a focus on transparency, fairness, accountability, privacy, and public trust. Extensive data were collected through a thorough review of purposefully selected written records, including books, journal articles, and book chapters. The gathered information was examined using qualitative thematic data analysis techniques by studying existing ethical frameworks, governance models, and successful practices. The paper suggests approaches and principles for encouraging the ethical use of AI in public service, considering the legal, ethical, and governance factors relevant to various situations. The paper seeks to offer valuable insights and strategies for effectively utilising artificial intelligence (AI) in public service delivery, to enhance the ethical deployment of citizen-centric AI and ultimately improve service delivery in public entities

Keywords: Artificial Intelligence (AI); Public Service Delivery; Implementation Strategies; Socio-Economic Challenges; Citizen-Centric; Data-Driven Decisions

1. Introduction

The rapid emergence and proliferation of digital technologies have reached an unprecedented level, enabling artificial intelligence artefacts to gain an extensive understanding of our lives and employ that knowledge to make decisions on our behalf, often in unforeseen ways. These advancements have the potential to either solve our most pressing issues or bring about catastrophic consequences, posing a dual risk of transforming our world for the better or leading to its destruction (Dignum, 2023; Stoica et al., 2017). Given these realities, artificial intelligence has emerged as a transformative force with the potential to revolutionise various sectors, including public service delivery (Di Vaio, Hassan & Alavoine, 2022). By utilising artificial intelligence technologies, public entities can bridge existing gaps, enhance service delivery outcomes, and establish a more inclusive and citizen-centric governance framework.

Nevertheless, effectively incorporating AI into public service delivery necessitates a thorough comprehension of potential risks, ethical considerations, and implementation approaches. The fundamental knowledge about the ethical strategies and guidelines for deploying AI in public service remains primarily rudimentary, conceptual and normative, offering insufficient evidence to better understand the potential risks, ethical implications and strategies of AI deployment in the public sector. This paper aims to explore potential risks, ethical implications, strategies, and guidelines for promoting ethical AI adoption in public service organisations. Stemming from the analyses of existing AI technology initiatives in public service, the paper provides insightful strategies and guidelines for successful AI integration, considering legal, ethical, and governance aspects while revealing the most significant potential risks and ethical implications. Two questions form the core of the paper:

1. What are the potential risks and societal implications associated with the adoption of AI in public service, including privacy, security, bias, and the impact on public trust in government services?

2. What governance mechanisms and regulatory frameworks can be implemented to ensure the responsible development and deployment of AI in public service, considering factors such as transparency, fairness, accountability, and public trust?

The rest of this paper is structured as follows: conceptualisation of AI technologies for public service delivery, research approach used in the study, results from the study and the proposed strategies and guidelines for deploying safe and ethical AI in public service.

1.2. Conceptualising and Contextualising the Utilisation of Artificial Intelligence Technologies in the Public Sector

Artificial intelligence in the public sector refers to the application of intelligent systems and algorithms to enhance government operations, improve service delivery, and support decision-making processes (Van Noordt & Misuraca, 2022; Valle-Cruz et al., 2020). It encompasses the utilisation of technologies like machine learning, natural language processing, computer vision, and robotics to automate tasks, analyse data, and derive insights.

AI, or artificial intelligence, is a branch of computer science that focuses on creating intelligent systems that can perform tasks that typically require human intelligence (Bellman, 1978; Sarker, 2022; Jarrahi, Askay, Eshraghi & Smith, 2023). Artificial intelligence involves the development and deployment of algorithms and techniques that enable machines to simulate and exhibit cognitive abilities, problem-solving skills, and decision-making capabilities (Bello et al., 2016; Luxton, 2014; Ameen et al., 2022).

The rapid advancement of artificial intelligence has the potential to revolutionise the public sector, enabling governments to improve service delivery, enhance decision-making processes, and optimise resource allocation (Wirtz, Weyerer & Geyer, 2018). By harnessing the power of AI, government agencies aim to address complex challenges, enhance citizen satisfaction, and drive positive socioeconomic outcomes. Here are a few case studies highlighting AI projects for service delivery in the public sector.

Automated Conversations for Government Interactions.

AI chatbots like ChatGPT, Microsoft Bing, and Google Bard are poised to fundamentally transform how governments gather, analyse, and utilise data for policy implementation and service delivery (Vander Walddt, 2022). In recent years, AI-enabled virtual agents and chatbots have been widely adopted by public entities to assist customers, service providers, and staff (Achat & Brown, 2019; Schoeman, Moore, Seedat & Chen, 2021). For instance, in South Africa, the city of Cape Town implemented a WhatsApp bot to improve service delivery, allowing residents to monitor service requests and receive updates through popular messaging platforms (City of Cape Town, 2021). Similarly, the Singaporean government utilises chatbots to provide instant updates during emergencies, like the COVID-19 pandemic, and to offer general support to its citizens. These examples demonstrate how AI chatbots are playing a significant role in enhancing the efficiency of public services and increasing citizen engagement.

Preventive Maintenance for Public Infrastructure.

The South African National Roads Agency (SANRAL) utilises artificial intelligence-driven predictive maintenance systems to enhance the management of road infrastructure (Nethamba & Grobbelaar, 2023). By examining historical data on road usage, weather patterns, and maintenance efforts, the system forecasts potential issues and deterioration, facilitating proactive maintenance. By adopting this strategy, the likelihood of road failures decreases, traffic disruptions are minimised, infrastructure durability is prolonged, and motorist safety is improved.

Experts emphasise the system's capability to aid maintenance managers in making well-informed choices, guaranteeing the safety of assets and the uninterrupted availability of the network. It also provides real-time asset health monitoring, allowing timely detection of issues, while remote inspections address problems in hazardous environments (Daniyan, Mpofo & Adeodu, 2020; Teoh, Gill & Parluka, 2021; Gbadamosi et al., 2020).

Globally, predictive maintenance is revolutionising the way infrastructure is managed. Japan's high-speed rail system employs IoT sensors and AI algorithms to monitor tracks and train parts, minimising delays and improving passenger security (Tsuruta et al., 2023). In Germany, cities like Hamburg employ IoT-enabled systems for water pipeline monitoring, analysing pressure and flow rates to detect leaks, improving efficiency, and lowering operational costs (Kataria, Rani & Kautish, 2024). These examples highlight the effectiveness of AI-powered predictive maintenance in enhancing the reliability, safety, and cost-effectiveness of infrastructure.

Analysing Healthcare Data for Tailored Treatment Plans.

The implementation of Electronic Health Records (EHR) is a significant catalyst for the growth of the healthcare analytics industry. In South Africa, healthcare providers increasingly use EHR systems to enhance patient care and reduce costs by capturing, storing, and sharing patients' medical information (Katurura & Cilliers, 2018; Sujansky & Kunz, 2015; Dubovitskaya et al., 2020).

Healthcare analytics is advancing personalised and precision medicine, which tailors treatments to individual patients based on genetics, lifestyle, history, and environment (Delpierre & Lefèvre, 2023; Subramanian et al., 2020).

By utilising data from EHRs and other sources, healthcare analytics enhances patient outcomes, detects patterns, and guides clinical choices, leading to cost reductions for healthcare providers (Garcelon et al., 2020; Sauer et al., 2022).

Globally, healthcare analytics applications demonstrate their capacity. The National Health Service (NHS) in the UK utilises artificial intelligence tools, such as DeepMind's Streams, to anticipate and prevent conditions like kidney failure, resulting in fewer hospital admissions and deaths (Topol, 2019). In India, advanced artificial intelligence systems like Netra AI address diabetes by developing customised care plans and optimising treatment strategies (Mackenzie et al., 2024; Sharma et al., 2024). Israeli companies like Medtronic and Tytocare incorporate analytics into wearable devices, allowing for personalised care for chronic conditions like hypertension and heart disease (Christodouleas et al., 2018; Gogia, 2019; Joshi, 2024).

Despite these advancements, ethical challenges persist, particularly in terms of transparency, fairness, privacy, and accountability. Addressing these issues is crucial for the responsible and equitable implementation of artificial intelligence (AI) technologies in healthcare (Shin & Park, 2019; Ahmad et al., 2022).

2. Methodology

The study primarily utilised secondary data that was purposefully and systematically selected, as it was deemed relevant and aligned with the research objectives. The step-by-step process of conducting a systematic literature review in this research is as follows: (1) identifying relevant studies, (2) selecting appropriate studies, (3) evaluating the relevance of the selected studies, and (4) synthesising the data obtained from the studies. A comprehensive account of each stage in the research process is provided below.

2.1 Step 1: Recognition of Research.

The first step was to establish the objectives and questions that would shape the literature review process. The objectives were established to situate the research within the existing body of knowledge and to provide valuable insights into the potential risks and ethical considerations associated with the use of artificial intelligence in public governance. To accomplish these goals, the following research questions were developed:

- a. What are the potential risks and societal implications associated with the adoption of AI in public governance?
- b. What governance mechanisms and regulatory frameworks can be put in place to ensure the responsible development and deployment of artificial intelligence in public service?

To answer these questions, the author utilised three different databases to locate scientific studies that were relevant to the topic: Web of Science, Scopus, and Google Scholar. These sources were chosen due to their extensive coverage of peer-reviewed literature and their direct relevance to the study's specific focus.

2.2 Step 2: Choice of Research.

The search process focused on journal articles and conference proceedings published between 2019 and 2024, written in English, and sourced from the Web of Science, Scopus, and the Digital Government References Library (DGRL). The scope was narrowed down by focusing on specific research areas to guarantee the relevance of the results. Scopus and Web of Science classify research fields in various ways. For Scopus, the search was restricted to four fields: social sciences, decision sciences, multidisciplinary, and business, management, and accounting. For Web of Science, the focus was on ten disciplines: public administration, library/information science, political science, management, communication, multidisciplinary sciences, engineering, international relations, and telecommunications.

The search strategy was executed in two stages. In April 2024, an initial search was performed to identify studies on the potential risks and implications of AI in public governance. This yielded 142 papers: 52 from Web of Science, 50 from Scopus, and 40 from the DGRL. For Scopus, only the top 30 most relevant results (as determined by the database) were included.

Given the increasing focus on AI research and its implications for public governance, the search was updated in January 2025 to capture recent developments. This yielded additional results: three new papers from Web of Science, four from Scopus, and one from the DGRL. To further ensure inclusivity, a supplementary search was conducted on google scholar, examining the first 50 results using keywords such as "AI," "artificial intelligence governance," "artificial intelligence policy," and "artificial intelligence government policy." however, this search did not yield new results, several papers already identified in web of science, Scopus, and the DGRL also appeared in google scholar. After eliminating and combining duplicate entries, the final dataset consisted of 75 distinct studies out of the initial 130 identified. This thorough and targeted approach guaranteed a comprehensive and relevant collection of literature for the study.

2.3 Step 3: Study Relevance and Quality Assessment

The third step in the systematic literature review process involved assessing the relevance and quality of the selected studies. This phase was divided into two key stages. First, the Author reviewed the titles and abstracts of the 75 identified studies to evaluate their relevance. Three criteria were applied during this initial screening:

- AI Focus: The study had to feature AI as a substantial or major component of its research questions or objectives. Studies where AI played a minor or secondary role were excluded.
- Public Governance Context: The study needed to address AI within the context of public governance (or the public sector). Articles that did not discuss AI in this context were removed.
- Core governance implications: the study had to focus on the implications of artificial intelligence use in public governance. Studies that briefly mentioned artificial intelligence or did not establish a connection to public governance were not included in the analysis.

After thoroughly examining the abstracts, 40 studies were discarded due to not meeting the specified criteria, leaving 35 studies to move forward to the next phase.

The second phase of the evaluation process required a comprehensive analysis of the 35 remaining studies. This comprehensive analysis assisted in identifying the ultimate choice. In this phase, nine studies were excluded as they did not adequately examine the intersection between artificial intelligence and public governance. In the end, 26 studies were chosen to be included in the review, as they directly focused on answering questions about the use of AI in public governance

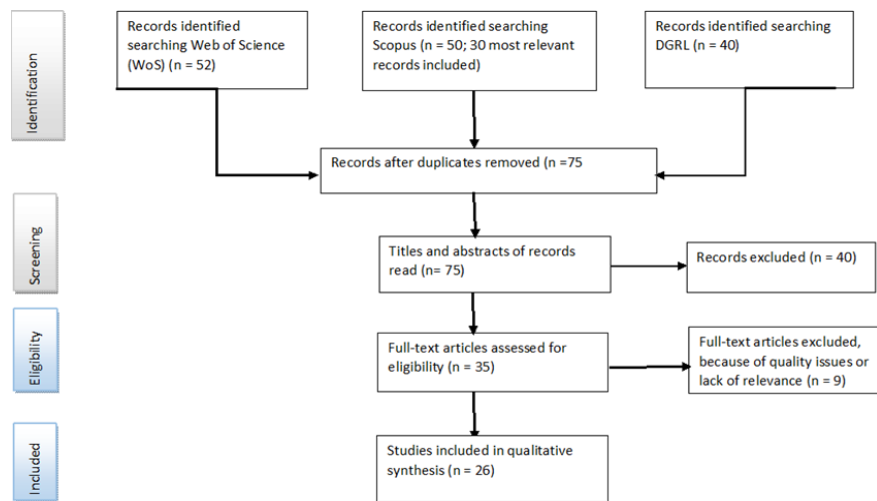


Figure 1. The Process of Selecting, Evaluating, and Including Studies for Analysis was Presented Using a PRISMA Flow Diagram.

2.4 Step 4: Data Examination.

The last stage of the research process included data examination. The information gathered during the literature review was carefully examined and organised, and the results are outlined in section 4. Thematic analysis was the main approach used to analyse the data.

3. Results and Discussion

This section presents the findings from the comprehensive analysis of the gathered documents. Specifically, it emphasises the key findings related to the potential risks and challenges of using artificial intelligence in government entities (3.1) and the moral and ethical considerations surrounding its use (3.2).

By combining the findings from the literature analysis (Cooper, 2017), the author has formulated a research agenda on utilising artificial intelligence in public governance (3.3).

3.1 Risks of Deploying AI in the Public Sector.

The incorporation of artificial intelligence in public service brings about various advantages, but it also entails potential risks that require thorough examination. Based on the comprehensive examination of the gathered documents, the study identifies the following significant risks linked to the utilisation of AI in public service.

Prejudice and Equity.

Bias in AI refers to the existence of systematic and unjustified inclinations or prejudices in the data utilised for training AI models or within the algorithms themselves. Such biases can arise from historical inequalities, human subjectivity, or unrepresentative datasets, leading to discriminatory outcomes that reinforce societal prejudices (Pulivarthy & Whig, 2025, Kausar et al., 2025). For example, biases can seep into artificial intelligence systems during the stages of data collection, preprocessing, and model training, leading to unfair treatment of specific demographic groups, including minorities, women, or economically disadvantaged individuals.

Nevertheless, the notion of fairness is of utmost importance in tackling and minimising these biases. Fairness in AI aims to guarantee that everyone is treated fairly, preventing any group from being unfairly disadvantaged by

AI-driven processes. It is essential for ensuring that outcomes do not favour or discriminate based on characteristics such as race, gender, ethnicity, or socioeconomic status (Albous & Alboushi, 2025; Decker, Wegner & Leicht-Scholten, 2025). By prioritising fairness, AI systems can be developed to prevent the systemic biases that may result in disparities, guaranteeing equitable opportunities in public service delivery and decision-making processes. The intersection between bias and fairness highlights the significance of recognising, resolving, and minimising discriminatory outcomes while ensuring equal treatment in artificial intelligence applications.

Openness and Responsibility

When evaluating the potential risks of using artificial intelligence in public service, it is crucial to prioritise transparency and accountability. Both play a crucial role in ensuring that artificial intelligence systems are deployed ethically, responsibly, and in the best interest of the public. Understanding and disclosing AI systems' decision-making processes is an arduous task. The absence of clear accountability or transparency in artificial intelligence models can lead to substantial legal risks and potential harm (Ashraf & Mustafa, 2025; Poli, Pamidi & Poli, 2025; Softech, 2025).

Transparency is crucial for establishing public trust, facilitating stakeholder comprehension, and enabling the examination of artificial intelligence systems to ensure they adhere to ethical and legal guidelines. Nevertheless, artificial intelligence algorithms are frequently intricate and challenging to comprehend, making it hard to grasp the reasoning behind the decisions they make. The absence of transparency can lead to accountability problems, making it difficult to hold artificial intelligence systems and their operators accountable for errors or biases, and eroding trust in their fairness and effectiveness.

Data Privacy Issues.

Safety concerns are of the utmost importance when evaluating the potential risks involved in the utilisation of artificial intelligence in public service. The implementation of artificial intelligence in multiple public sectors, such as healthcare, law enforcement, transportation, and disaster response, presents a range of safety-related challenges. The study identified several key safety concerns associated with the use of AI in public service, including:
Poor Judgment.

AI systems can sometimes make mistakes or inaccurate predictions because their algorithms are not perfect or they haven't been trained with enough data. In critical applications like healthcare diagnosis or law enforcement, errors in public service can have grave repercussions, jeopardising public safety and undermining trust. Furthermore, the quality of data utilised by artificial intelligence systems is frequently disregarded. Incomplete, inconsistent, or incorrect data can result in flawed analysis and harmful choices. These flawed decisions could negatively affect the organisation and the public (Alifia, Safitri, Irhami & Kusumasari, 2025; Reddy, Reddy, Pilly & doss, 2025; Srinivasan, Abd Rahman & Ravana, 2025).

Protection of Personal and Confidential Information

Public service applications frequently handle confidential, individual data. Inadequate security measures can result in data breaches, posing risks to privacy and public safety. Unauthorised access to this information can result in identity theft, blackmail, or other forms of personal harm. The rise of cyber threats and sophisticated attacks highlights the need for robust security protocols to safeguard AI systems in the public sector. AI-driven defensive systems, too, are vulnerable to attacks that exploit inherent flaws (Akinade, Adepoku, Ige & Afolabi, 2025; Anju, Freeda & Venket, 2025; Damodaran, Edwin, Ramathilagam, Jeno & Suganya, 2025; Kiranbabu et al., 2025).

Threats and Breaches of Data

As artificial intelligence systems become increasingly important in critical sectors, their vulnerability to cyberattacks raises substantial concerns. These attacks can lead to unauthorised access to sensitive data, privacy violations, or even manipulation of artificial intelligence functionalities. For instance, attackers could exploit weaknesses in AI algorithms to alter decision-making processes, disrupting services like healthcare, transportation, or utilities (Chakraborty, Biswas & Khan, 2023; Guembe et al., 2024; Prince et al., 2024; Yamin et al., 2021).

The repercussions of these breaches extend beyond mere technical disruptions, encompassing ethical and societal ramifications. Unauthorised access to data can infringe upon privacy rights, erode public trust, and subject individuals to potential harm. Furthermore, compromised AI systems may facilitate large-scale fraud, spread misinformation, or endanger public safety through the malfunctioning of critical infrastructure.

To mitigate these risks, it is crucial to implement strong security measures, including encryption, constant monitoring, and regular vulnerability assessments. Furthermore, ethical governance frameworks need to be established to guarantee responsible AI design and deployment, reducing the risk of harm and prioritising the well-being of the public.

Ensuring Fairness in Access to AI.

The widespread adoption of AI in public service has the potential to reshape society in ways that are difficult to fully anticipate (Medaglia, Gil-Garcia & Pardo, 2023; Wilson & van der Velden, 2022). Ethical considerations

encompass alterations in cultural norms, shifts in power dynamics, and the reconfiguration of social hierarchies. For example, the automation of decision-making processes may decrease human oversight, which could lead to a reduction in accountability and transparency in public administration.

Moreover, AI-driven systems could exacerbate existing inequalities if marginalised groups face limited access to technology or if biases in AI algorithms perpetuate discrimination (Oladoyinbo et al., 2024; Roche, Wall & Lewis, 2023). The increasing dependence on AI could impact the job market, potentially causing job losses or the creation of new positions that demand advanced technical expertise, thereby reshaping employment patterns and social mobility.

These alterations bring up important concerns about the lasting effects of AI on society, particularly in terms of fairness, inclusivity, and overall welfare. Overcoming these challenges necessitates a proactive approach, which involves actively involving the public, implementing ethical oversight mechanisms, and establishing policies that prioritise the well-being of society while mitigating potential risks.

The Impact of Automation on Employment.

The potential for job displacement is a significant worry when evaluating the risks involved in implementing AI in public service. The incorporation of artificial intelligence (AI) in different public sector roles can have a substantial impact on employment trends. Studies on the influence of artificial intelligence on the job market have produced a variety of outcomes. Wang and Lu (2025) highlight the varied results, with AI potentially boosting economic growth but also altering the job market. Bruun and Duka (2018) note that while AI enhances productivity, it has also led to job displacement, sparking conversations about potential policy measures like universal basic income to mitigate social instability.

Conversely, Jarrahi (2018) and Nguyen and Vo (2022) present a more positive perspective, emphasising AI's capacity to enhance decision-making and enhance job quality. Song et al. (2022) contend that the integration of artificial intelligence with human labour can lead to superior results. Kaplan and Haenlein (2020) and Mutascu (2021) stress that AI primarily replaces low-skilled jobs, reshaping the labour market's composition.

Beyond physical modifications, AI also influences mental facets. Verma and Singh (2022) and Li et al. (2023) emphasise how fears of becoming outdated and decreased motivation can disproportionately impact specific groups, intensifying societal worries. More recent research by Wang, Weng, and Lu (2025) finds that highly educated, female, older, and remote workers perceive greater risks of job displacement due to AI, owing to factors like digital divides and AI's increasing relevance to their fields. Despite these concerns, the broader consensus suggests that AI's substitution risks are most prominent in low-skilled, repetitive jobs (Decanio, 2016, Frey & Osborne, 2017, Verma & Singh, 2022; Dwivedi et al., 2021).

Compliance with Laws and Regulations.

The legal and regulatory framework governing privacy in the context of artificial intelligence is frequently lacking or not uniformly enforced. Insufficient legislation and weak enforcement mechanisms create gaps in protecting individuals' privacy, highlighting the need for comprehensive, up-to-date regulations to address emerging challenges (Elbaih, 2023; Rodrigues, 2020; Katyral, 2019). Additionally, the fast-paced advancement of artificial intelligence technology surpasses the legislative process, posing challenges for current laws to adequately address emerging privacy concerns. In the absence of adaptable and responsive regulatory frameworks, antiquated laws may prove inadequate in protecting individuals' rights amidst the rapid advancements of artificial intelligence. The lack of uniform enforcement across different jurisdictions worsens the issue, leading to variations in privacy safeguards and potentially eroding public confidence in artificial intelligence systems utilised in critical sectors such as healthcare, education, and law enforcement.

3.2 Ethical and Moral Considerations of AI in Public Service.

This section of the study presents discussions on the emerging themes on the moral consequences of AI in the public service

Prejudice and Equity

The ethical considerations surrounding artificial intelligence, particularly concerning bias and fairness, are of utmost importance in ensuring its responsible integration (Emma, 2023; Modi, 2024). AI systems, influenced by the data and algorithms they depend on, can inadvertently reinforce or amplify societal biases, resulting in unfair outcomes that disproportionately impact marginalised groups (Bose, 2025; Dinker, 2024; Lainjo, 2020). These challenges raise concerns about openness, responsibility, and fairness, underscoring the importance of ethical guidelines to shape the development and use of artificial intelligence. Ensuring fairness and eliminating bias is crucial for building trust and equity in AI-driven solutions, particularly in public service settings where decisions can have a profound impact on individuals and communities.

Openness and Responsibility

The ethical considerations surrounding the adoption of artificial intelligence revolve around transparency and accountability. The opaque nature of many AI systems makes it difficult to trace how decisions are made, creating challenges in holding individuals or organisations responsible for errors or unfair outcomes (Cheng, Varshney & Liu, 2021; Lepri et al., 2018; Scatiggio, 2020). In fields such as public service, where choices can have significant impacts on society, the absence of clarity erodes trust and fairness. Creating transparent and easily comprehensible procedures and holding individuals responsible are crucial for constructing systems that are both efficient and morally upright.

Safeguarding Personal Information.

The ethical considerations surrounding privacy and data protection in the field of artificial intelligence are of utmost importance. AI systems heavily rely on data, but this dependence exposes them to potential misuse, breaches, and the erosion of individual privacy. Without proper safeguards, personal information can be compromised or misused, eroding public trust in the system. In the realm of public service, where confidential information is frequently encountered, prioritising privacy and implementing stringent data protection measures is crucial to maintain ethical standards and ensure responsible AI deployment.

The extensive utilisation of personal data in AI raises concerns about data security, misuse, and potential breaches, particularly in public services dealing with sensitive information.

Human Control and Machine Independence.

Autonomy and human oversight are crucial ethical considerations when implementing artificial intelligence. While AI can automate decision-making processes, excessive dependence on machines can diminish human autonomy and critical thinking. Without proper human supervision, artificial intelligence systems may make mistakes or produce outcomes that lack the sensitivity and understanding required in intricate situations (Jedličková, 2024; Konidena, Malaiyappan & Tadimari, 2024). In the realm of public service, where choices impact individuals' lives, it is crucial to prioritise human control and accountability to uphold ethical principles and prevent unforeseen outcomes.

The excessive dependence on AI could diminish human decision-making roles and undermine accountability (Spatola, 2024), resulting in ethical dilemmas when mistakes happen. Ensuring human supervision is crucial in order to address these concerns.

Public Confidence and Social Approval.

Public trust and social acceptance are critical pillars for the ethical deployment of AI, particularly in the public sector. The effectiveness of AI technologies hinges not only on their technical performance but also on the public's perception of their fairness, transparency, and alignment with societal values. Inadequate attention to ethical concerns can lead to a breakdown in trust (Adelakun, 2024; Dirgová Luptáková, Pospíchal & Hhuraj, 2023; Wang & Wu, 2024) with the public questioning the motives behind AI-driven decisions. When AI is perceived as opaque, biased, or unaccountable, it can face significant resistance, ultimately hindering its potential to drive positive change. Therefore, cultivating trust through transparent practices, ethical guidelines, and open dialogue is essential to ensuring that AI remains a trusted tool for societal progress. For public service to be successful, citizens must have faith in its fairness and dependability. Misuse of AI can lead to a loss of trust, which can hinder its adoption and effectiveness.

Equality and Availability.

Equity and access are essential ethical concerns when implementing AI, particularly in the public service sector (Mikhaylov, Esteve & Campion, 2018; Trajkovski, 2024). As AI technologies become more widespread, there is a risk that marginalised groups may be left behind due to unequal access to the necessary resources, such as digital infrastructure and technological literacy (Bach, Shaffer & Wolfson, 2013; Helsper, 2021; Olanrewaju, Adebayo, Omotosho & Olajide, 2021). The digital divide can worsen existing social disparities, hindering certain groups from reaping the advantages of AI-powered services. To ensure that AI contributes to social good, it is crucial to design systems that are inclusive, accessible, and equitable, providing equal opportunities for all individuals, regardless of their socio-economic status or geographic location. The digital divide can worsen disparities, as not all communities have equal access to the technologies needed to benefit from AI-driven public services.

3.3 Agenda on Strategies for Utilising Artificial Intelligence in Public Service

The successful implementation of artificial intelligence in the public sector requires thoughtful, careful consideration of several key components (Susar & Aquaro, 2019; Zafar & Villeneuve, 2018; Wirtz, Weyerer & Geyer, 2019). These elements serve as the basis for harnessing the transformative power of AI while tackling ethical, technical, and organisational obstacles. Let's examine these crucial elements.

Data governance and Control.

Data governance and management are critical components of ai in the public sector (Mäntymäki, Minkinen, Birkstedt & Viljanen, 2022, Zuiderwijk, Chen & Salem, (2021). Governments must establish robust frameworks for

data collection, storage, quality assurance, privacy protection, and security (Rizi & Seno, 2022; Janssen, Brous, Estevez, Barbosa, & Janowski, 2020). This involves defining data ownership, access rights, and data-sharing protocols (Filgueiras & Lui, 2023; Zorrilla & Yebenes, 2022; Mahanti & Mahanti, 2021). Effective data governance ensures that AI systems are fuelled by accurate, reliable, and diverse datasets from various sources, enabling better decision-making and policy formulation (Kuziemski & Misuraca, 2020; Ladley, 2019; Ren, Zhang, Liu et al., 2019).

Infrastructure and Technical Readiness

The successful implementation of AI in the public sector necessitates the right technical infrastructure and capabilities (Desouza, Dawson, & Chenok, 2020; Susar & Aquaro, 2019; Wirtz, Weyerer, Geyer, 2019; Mikalef et al., 2022). Governments must secure scalable computing resources, high-speed connectivity, and cloud-based services to meet the computational demands of AI algorithms. Notably, the specific infrastructure needs for AI can vary depending on the scale, complexity, and objectives of the AI system being developed. Governments must invest in building technical expertise within their agencies (Susar & Aquaro, 2019; Mikhaylov, Esteve, & Champion, 2018) and foster partnerships with AI experts (Aerts & Bogdan-Martin, 2021, Kayembe & Nel, 2019, Schmidt et al., 2021), while leveraging external resources to develop and maintain AI systems effectively (Desouza, Dawson, & Chenok, 2020; Mikhaylov, Esteve, & Champion, 2018).

Legal and Regulatory Frameworks

The development of legal and regulatory frameworks is crucial for responsible AI deployment in the public sector. Governments must establish laws and regulations that govern data protection (Anagnostou et al., 2022; Sawhney & Torres, 2022), algorithmic accountability (Busuioc, 2021; Henman, 2020), privacy, and ethical considerations (Chakrabarti & Sanyal, 2020; Golbin, Rao, Hadjarian & Krittman, 2020; Champion, Gasco-Hernandez, Jankin Mikhaylov, & Esteve, 2022). These frameworks address issues such as algorithmic bias, transparency in decision-making, accountability of AI systems, and the protection of individual rights. They help ensure that AI technologies align with societal values, uphold fairness, and mitigate potential risks.

Development of Skills and Knowledge.

Developing literacy and skills in the public sector workforce is crucial for achieving successful AI implementation. Governments need to invest in training programmes and capacity-building initiatives to equip employees with the necessary knowledge and skills to work effectively with AI technologies. This includes training in data analytics, machine learning, and artificial intelligence ethics. Collaboration and knowledge sharing among government agencies, academic institutions, and industry experts are essential for fostering a culture of continuous learning and innovation.

Stakeholder Engagement and Collaboration.

AI implementation in the public sector should involve active stakeholder engagement and collaboration (Mikhaylov, Esteve, & Champion, 2018) and foster partnerships with AI experts (Aerts & Bogdan-Martin, 2021; Kayembe & Nel, 2019; Schmidt et al., 2021). Governments should seek input from citizens, civil society organisations, academia, and industry to ensure inclusive and transparent decision-making processes (Kuziemski & Misuraca, 2020; Valle-Cruz, Criado, Sandoval Almazan & Ruvalcaba-Gomez, 2020). Engaging stakeholders helps in understanding their needs, addressing concerns, and building trust. Collaboration with AI experts and industry stakeholders facilitates knowledge exchange, access to cutting-edge technologies, and shared learning experiences (Sen, 2023; Malik, de Silva, Budhwar, & Srikanth, 2021; Olan, Arakpogun, Sulkal, Nakpodia, Damij & Jayawickrama, 2022).

Analysis and Significance of Our Outcome.

Continuous evaluation and impact assessment are critical for monitoring the effectiveness, efficiency, and societal impact of AI in the public sector (Zuidervijk, Chen & Salem, 2021, Valle-Cruz, Criado, Sandoval Almazan & Ruvalcaba-Gómez, 2020). Governments should establish evaluation frameworks and performance metrics to assess the outcomes and value generated by AI systems (Kuziemski & Misuraca, 2020; König & Wenzelburger, 2021). This involves evaluating the accuracy, fairness, and transparency of AI algorithms, as well as assessing the broader impact on public services, citizen engagement, and social equity (Schiff, Schiff & Pierson, 2022; Keddell, 2019). Feedback mechanisms, user surveys, and citizen participation are essential in identifying areas for enhancement and maintaining accountability.

By prioritising these essential elements, governments can successfully navigate the intricacies of implementing AI in the public sector. Addressing data governance, technical readiness, legal frameworks, capacity building, stakeholder engagement, and evaluation facilitates responsible and impactful ai deployment, leading to improved public services, evidence-based decision-making, and enhanced societal outcomes.

4. Conclusion and Recommendations

This research aimed to investigate the potential risks and ethical concerns related to the integration of artificial intelligence (AI) in public service, with a specific emphasis on transparency, fairness, accountability, privacy, and public trust. It has offered valuable insights into the societal consequences of integrating AI into this context, encompassing concerns about data privacy, security, algorithmic bias, and the potential erosion of public trust in government institutions. Recognising that the responsible development and deployment of AI in the public sector is a matter of broad societal concern, the paper offers guidance on governance mechanisms and regulatory frameworks that could support ethical and accountable use of AI, ultimately aimed at enhancing public service delivery.

This study aims to shed light on the ethical dilemmas associated with the integration of AI into public services and offers practical solutions to mitigate these challenges, thereby enriching the existing body of knowledge on responsible AI adoption. It is crucial to have a comprehensive understanding of these risks and ethical considerations to ensure that AI is implemented in a manner that enhances service delivery while minimising unintended consequences.

Further Study

Future studies and policy initiatives should build upon these findings to ensure that artificial intelligence is utilised ethically and efficiently for the benefit of society.

References

- [1] Adelakun, D. (2024). Organizational factors that influence fairness in algorithmic decision-support.
- [2] Aerts, A., & Bogdan-Martin, D. (2021). Leveraging data and AI to deliver on the promise of digital health. *International Journal of Medical Informatics*, 150, 104456.
- [3] Ahmad, K., Maabreh, M., Ghaly, M., Khan, K., Qadir, J., & Al-Fuqaha, A. (2022). Developing future human-centred smart cities: Critical analysis of smart city security, Data management, and Ethical challenges. *Computer Science Review*, 43, 100452.
- [4] Akinade, A. O., Adepoju, P. A., Ige, A. B., & Afolabi, A. I. (2025). Cloud Security Challenges and Solutions: A Review of Current Best Practices.
- [5] Akinade, Adepoju, Ige & Afolabi, 2025; Anju, Freeda & Venket, 2025; Damodaran, Edwin, Ramathilagam, Jeny & Suganya, 2025; Kiranbabu et al., 2025).
- [6] Albous, M. R., & Alboloushi, B. (2025). AI-Driven Innovations in E-Government: How Is AI Reshaping the Public Sector?. In *Hamessing AI, Blockchain, and Cloud Computing for Enhanced e-Government Services* (pp. 93-118). IGI Global Scientific Publishing.
- [7] Alifia, R. A., Safitri, N. R., Irhami, D. M., & Kusumasari, I. R. (2025). Challenges and Solutions for Decision Making in the Era of Big Data. *Jurnal Bisnis dan Komunikasi Digital*, 2(2), 13-13.
- [8] Alvarenga, A., Matos, F., Godina, R., & CO Matias, J. (2020). Digital transformation and knowledge management in the public sector. *Sustainability*, 12(14), 5824.
- [9] Ameen, N., Sharma, G. D., Tarba, S., Rao, A., & Chopra, R. (2022). Toward advancing theory on creativity in marketing and artificial intelligence. *Psychology & Marketing*, 39(9), 1802-1825.
- [10] Anagnostou, M., Karvounidou, O., Katritzidaki, C., Kechagia, C., Melidou, K., Mpeza, E., ... & Peristeras, V. (2022). Characteristics and challenges in the industries towards responsible AI: a systematic literature review. *Ethics and Information Technology*, 24(3), 37.
- [11] *Analysis and Policy*, 69, 653–667. <https://doi.org/10.1016/j.eap.2021.01.012>
- [12] Anju, A., Freeda, A. R., & Venket, K. (2025). Security Concerns With ChatGPT and Other AI Tools. In *Real-Time Data Decisions With AI and ChatGPT Techniques* (pp. 221-236). IGI Global.
- [13] Ashraf, Z. A., & Mustafa, N. (2025). AI Standards and Regulations. *Intersection of Human Rights and AI in Healthcare*, 325-352.
- [14] Bach, A., Shaffer, G., & Wolfson, T. (2013). Digital human capital: Developing a framework for understanding the economic impact of digital exclusion in low-income communities. *Journal of Information Policy*, 3, 247-266.
- [15] Bellman R (1978) *An introduction to artificial intelligence: can computers think?*. Boyd & Fraser, San Francisco
- [16] Bello, O., Teodoru, C., Yaqoob, T., Oppelt, J., Holzmann, J., & Obiwanne, A. (2016). Application of artificial intelligence techniques in drilling system design and operations: a state of the art review and future research pathways. In *SPE Nigeria Annual International Conference and Exhibition*. OnePetro.
- [17] Bose, M. (2025). Bias in AI: A Societal Threat: A Look Beyond the Tech. In *Open AI and Computational Intelligence for Society 5.0* (pp. 197-224). IGI Global Scientific Publishing.

- [18] Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review*, 81(5), 825-836.
- [19] Champion, A., Gasco-Hernandez, M., Jankin Mikhaylov, S., & Esteve, M. (2022). Overcoming the challenges of collaboratively adopting artificial intelligence in the public sector. *Social Science Computer Review*, 40(2), 462-477.
- [20] Chakrabarti, R., & Sanyal, K. (2020). Towards a 'Responsible AI': can India take the lead?. *South Asia Economic Journal*, 21(1), 158-177.
- [21] Chakraborty, A., Biswas, A., & Khan, A. K. (2023). Artificial intelligence for cybersecurity: Threats, attacks and mitigation. In *Artificial Intelligence for Societal Issues* (pp. 3-25). Cham: Springer International Publishing.
- [22] Cheng, L., Varshney, K. R., & Liu, H. (2021). Socially responsible ai algorithms: Issues, purposes, and challenges. *Journal of Artificial Intelligence Research*, 71, 1137-1181.
- [23] Christodouleas, D. C., Kaur, B., & Chorti, P. (2018). From point-of-care testing to eHealth diagnostic devices (eDiagnostics). *ACS Central Science*, 4(12), 1600-1616.
- [24] Cooper, H. (2017). *Research synthesis and meta-analysis : a step-by-step approach* (5th ed., Vol. 2).
- [25] Damodaran, D., Edwin, R. S., Ramathilagam, A., Jenö, J. J., & Suganya, S. (2025). Ensuring Privacy and Security of EHR With Secure Collaborative Transfer Learning (SCTL). In *Generative AI Techniques for Sustainability in Healthcare Security* (pp. 213-242). IGI Global Scientific Publishing.
- [26] Decker, M. C., Wegner, L., & Leicht-Scholten, C. (2025). Procedural fairness in algorithmic decision-making: the role of public engagement. *Ethics and Information Technology*, 27(1), 1.
- [27] Desouza, K. C., Dawson, G. S., & Chenok, D. (2020). Designing, developing, and deploying artificial intelligence systems: Lessons from and for the public sector. *Business Horizons*, 63(2), 205-213.
- [28] Dhirani, L. L., Mukhtiar, N., Chowdhry, B. S., & Newe, T. (2023). Ethical dilemmas and privacy issues in emerging technologies: A review. *Sensors*, 23(3), 1151.
- [29] Di Vaio, A., Hassan, R., & Alavoine, C. (2022). Data intelligence and analytics: A bibliometric analysis of human–Artificial intelligence in public sector decision-making effectiveness. *Technological Forecasting and Social Change*, 174, 121201.
- [30] Dignum, V. (2023). Responsible Artificial Intelligence: Recommendations and Lessons Learned. In *Responsible AI in Africa: Challenges and Opportunities* (pp. 195-214). Cham: Springer International Publishing.
- [31] Dinker, N. (2024). Artificial Intelligence and Inequality: Examining the Social Divides Created by Technological Advancements. *International Journal of Innovations in Science, Engineering And Management*, 228-236.
- [32] Dirgová Luptáková, I., Pospíchal, J., & Huraj, L. (2023). Beyond Code and Algorithms: Navigating Ethical Complexities in Artificial Intelligence. In *Proceedings of the Computational Methods in Systems and Software* (pp. 316-332). Cham: Springer Nature Switzerland.
- [33] ElBaih, M. (2023). The role of privacy regulations in ai development (A Discussion of the Ways in Which Privacy Regulations Can Shape the Development of AI). Available at SSRN 4589207.
- [34] Emma, L. (2024). The Ethical Implications of Artificial Intelligence: A Deep Dive into Bias, Fairness, and Transparency.
- [35] Farahani, M., & Ghasemi, G. (2024). Artificial intelligence and inequality: Challenges and opportunities. *Int. J. Innov. Educ*, 9, 78-99.
- [36] Filgueiras, F., & Lui, L. (2023). Designing data governance in Brazil: an institutional analysis. *Policy Design and Practice*, 6(1), 41-56.
- [37] Gogia, S. (Ed.). (2019). *Fundamentals of telemedicine and telehealth*. Academic Press.
- [38] Golbin, I., Rao, A. S., Hadjarian, A., & Krittman, D. (2020). Responsible AI: a primer for the legal community. In *2020 IEEE International Conference on Big Data (Big Data)* (pp. 2121-2126). IEEE.
- [39] Guembe, B., Azeta, A., Misra, S., Osamor, V. C., Fernandez-Sanz, L., & Pospelova, V. (2022). The emerging threat of ai-driven cyber attacks: A review. *Applied Artificial Intelligence*, 36(1), 2037254
- [40] Helsper, E. (2021). The digital disconnect: The social causes and consequences of digital inequalities.
- [41] Henman, P. (2020). Improving public services using artificial intelligence: possibilities, pitfalls, governance. *Asia Pacific Journal of Public Administration*, 42(4), 209-221.
- [42] Ijaiya, H., & Odumuwan, O. O. Advancing Artificial Intelligence and Safeguarding Data Privacy: A Comparative Study of EU and US Regulatory Frameworks Amid Emerging Cyber Threats.
- [43] 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.
- [44] Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577–586. <https://doi.org/10.1016/j.bushor.2018.03.007>

- [45] Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons*, 66(1), 87-99.
- [46] Jedličková, A. (2024). Ethical approaches in designing autonomous and intelligent systems: a comprehensive survey towards responsible development. *AI & SOCIETY*, 1-14.
- [47] Joshi, N. (2024). Emerging Challenges in Privacy Protection with Advancements in Artificial Intelligence. *International Journal of Law and Policy*, 2(4), 55-77.
- [48] Joshi, P. (2024). Communication Protocols for Internet of Medical Things (IoMT). In *Interdisciplinary Technological Advancements in Smart Cities* (pp. 285-310). Cham: Springer Nature Switzerland.
- [49] Kaplan, A., & Haenlein, M. (2020). Rulers of the world, unite! The challenges and
- [50] Kataria, A., Rani, S., & Kautish, S. (2024). Artificial Intelligence of Things for Sustainable Development of Smart City Infrastructures. In *Digital Technologies to Implement the UN Sustainable Development Goals* (pp. 187-213). Cham: Springer Nature Switzerland.
- [51] Katyal, S. K. (2019). Private accountability in the age of artificial intelligence. *UCLA L. Rev.*, 66, 54.
- [52] Kausar, A., Kausar, A., Kulkarni, S., Bhosale, P. A., & Das, S. (2025). Gender Bias in AI: Creating Discriminatory Systems. In *Dimensions of Diversity, Equity, Inclusion, and Belonging in Business* (pp. 87-106). IGI Global Scientific Publishing.
- [53] Kayembe, C., & Nel, D. (2019). Challenges and opportunities for education in the Fourth Industrial Revolution. *African Journal of Public Affairs*, 11(3), 79-94.
- [54] Keddell, E. (2019). Algorithmic justice in child protection: Statistical fairness, social justice and the implications for practice. *Social Sciences*, 8(10), 281.
- [55] Kiranbabu, M. N. V., Viji, A. J., Chandanan, A. K., Birchha, V., Pandey, T. K., & Sar, S. K. (2025). The Challenge of Adversarial Attacks on AI-Driven Cybersecurity Systems. *Journal of Cybersecurity & Information Management*, 15(1).
- [56] Kitchenham, B. (2004). Procedures for performing systematic reviews. Keele, UK, Keele University, 33(2004), 1-26.
- [57] Koekemoer, S., & Von Solms, R. (2021). Does the 4th industrial revolution provide solutions for sustainable basic service delivery in South African municipalities? *Journal of Public Administration*, 56(2), 339-351.
- [58] Konidena, B. K., Malaiyappan, J. N. A., & Tadimarri, A. (2024). Ethical Considerations in the Development and Deployment of AI Systems. *European Journal of Technology*, 8(2), 41-53.
- [59] König, P. D., & Wenzelburger, G. (2021). The legitimacy gap of algorithmic decision-making in the public sector: Why it arises and how to address it. *Technology in Society*, 67, 101688.
- [60] 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.
- [61] Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications policy*, 44(6), 101976.
- [62] Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications policy*,
- [63] Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications policy*, 44(6), 101976.
- [64] Ladley, J. (2019). *Data governance: How to design, deploy, and sustain an effective data governance program*. Academic Press.
- [65] Lainjo, B. (2020). The global social dynamics and inequalities of artificial intelligence. *Int. J. Innov. Sci. Res. Rev*, 5, 4966-4974.
- [66] Latupeirissa, J. J. P., Dewi, N. L. Y., Prayana, I. K. R., Srikandi, M. B., Ramadiansyah, S. A., & Pramana, I. B. G. A. Y. (2024). Transforming public service delivery: A comprehensive review of digitization initiatives. *Sustainability*, 16(7), 2818.
- [67] Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes: The premise, the proposed solutions, and the open challenges. *Philosophy & Technology*, 31(4), 611-627.
- [68] Li, C., Zhang, Y., Niu, X., Chen, F., & Zhou, H. (2023). Does artificial intelligence promote
- [69] Luxton, D. D. (2014). Artificial intelligence in psychological practice: Current and future applications and implications. *Professional Psychology: Research and Practice*, 45(5), 332.
- [70] Mackenzie, S. C., Sainsbury, C. A., & Wake, D. J. (2024). Diabetes and artificial intelligence beyond the closed loop: a review of the landscape, promise and challenges. *Diabetologia*, 67(2), 223-235.

- [71] Mahanti, R., & Mahanti, R. (2021). Data Governance and Data Management Functions and Initiatives. *Data Governance and Data Management: Contextualizing Data Governance Drivers, Technologies, and Tools*, 83-143.
- [72] Malik, A., De Silva, M. T., Budhwar, P., & Srikanth, N. R. (2021). Elevating talents' experience through innovative artificial intelligence-mediated knowledge sharing: Evidence from an IT-multinational enterprise. *Journal of International Management*, 27(4), 100871.
- [73] Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics*, 2(4), 603-609.
- [74] Medaglia, R., Gil-Garcia, J. R., & Pardo, T. A. (2023). Artificial intelligence in government: Taking stock and moving forward. *Social Science Computer Review*, 41(1), 123-140.
- [75] Mihiyawi, S. (2024). The Artificial Intelligence Era Between Governance and Our Privacy Protection. Sameer Mihiyawi.
- [76] Mikalef, P., Lemmer, K., Schaefer, C., Ylinen, M., Fjørtoft, S. O., Torvatn, H. Y., ... & Niehaves, B. (2022). Enabling AI capabilities in government agencies: A study of determinants for European municipalities. *Government Information Quarterly*, 39(4), 101596.
- [77] Mikhaylov, S. J., Esteve, M., & Campion, A. (2018). Artificial intelligence for the public sector: opportunities and challenges of cross-sector collaboration. *Philosophical transactions of the royal society a: mathematical, physical and engineering sciences*, 376(2128), 20170357.
- [78] Mikhaylov, S. J., Esteve, M., & Campion, A. (2018). Artificial intelligence for the public sector: opportunities and challenges of cross-sector collaboration. *Philosophical transactions of the royal society a: mathematical, physical and engineering sciences*, 376(2128), 20170357.
- [79] Modi, T. B. (2023). Artificial Intelligence Ethics and Fairness: A study to address bias and fairness issues in AI systems, and the ethical implications of AI applications. *Revista Review Index Journal of Multidisciplinary*, 3(2), 24-35.
- [80] Mutascu, M. (2021). Artificial intelligence and unemployment: New insights. *Economic*
- [81] Nguyen, Q. P., & Vo, D. H. (2022). Artificial intelligence and unemployment: An international evidence. *Structural Change and Economic Dynamics*, 63, 40–55. <https://doi.org/10.1016/j.strueco.2022.09.003>
- [82] OECD (2020). *Artificial Intelligence in the Public Sector: Maximizing Opportunities, Managing Risks*. Paris: OECD Publishing.
- [83] Oladoyinbo, T. O., Olabanji, S. O., Olaniyi, O. O., Adebisi, O. O., Okunleye, O. J., & Ismaila Alao, A. (2024). Exploring the challenges of artificial intelligence in data integrity and its influence on social dynamics. *Asian Journal of Advanced Research and Reports*, 18(2), 1-23.
- [84] Olan, F., Arakpogun, E. O., Suklan, J., Nakpodia, F., Damij, N., & Jayawickrama, U. (2022). Artificial intelligence and knowledge sharing: Contributing factors to organizational performance. *Journal of Business Research*, 145, 605-615.
- [85] Olanrewaju, G. S., Adebayo, S. B., Omotosho, A. Y., & Olajide, C. F. (2021). Left behind? The effects of digital gaps on e-learning in rural secondary schools and remote communities across Nigeria during the COVID19 pandemic. *International journal of educational research open*, 2, 100092.
- [86] opportunities of artificial intelligence. *Business Horizons*, 63(1), 37–50. <https://doi.org/10.1016/j.bushor.2019.09.003>
- [87] org/10.1016/j.bushor.2019.09.003
- [88] Poli, P. K. R., Pamidi, S., & Poli, S. K. R. (2025). Unraveling the Ethical Conundrum of Artificial Intelligence: A Synthesis of Literature and Case Studies. *Augmented Human Research*, 10(1), 2.
- [89] Prince, N. U., Faheem, M. A., Khan, O. U., Hossain, K., Alkhayyat, A., Hamdache, A., & Elmouki, I. (2024). AI-Powered Data-Driven Cybersecurity Techniques: Boosting Threat Identification and Reaction. *Nanotechnology Perceptions*, 20, 332-353.
- [90] Pulivarthy, P., & Whig, P. (2025). Bias and Fairness Addressing Discrimination in AI Systems. In *Ethical Dimensions of AI Development* (pp. 103-126). IGI Global.
- [91] Puplampu, R. (2024). *What Everyone Should Know About the Rise of AI: AI Transparency, Privacy, and Ethics Best Practices*. Puplampu Books.
- [92] Reddy, C. K. K., Reddy, P. S., Pilly, A., & Doss, S. (2025). Transformative Effects of Smarter Chatbots: Unravelling the Vision, Challenges, and Capabilities of ChatGPT-Conversational AI. *Artificial Intelligence-Enabled Businesses: How to Develop Strategies for Innovation*, 333-350.
- [93] Ren, S., Zhang, Y., Liu, Y., Sakao, T., Huisingh, D., & Almeida, C. M. (2019). A comprehensive review of big data analytics throughout product lifecycle to support sustainable smart manufacturing: A framework, challenges and future research directions. *Journal of cleaner production*, 210, 1343-1365.
- [94] Rizi, M. H. P., & Seno, S. A. H. (2022). A systematic review of technologies and solutions to improve security and privacy protection of citizens in the smart city. *Internet of Things*, 100584.

- [95] Roche, C., Wall, P. J., & Lewis, D. (2023). Ethics and diversity in artificial intelligence policies, strategies and initiatives. *AI and Ethics*, 3(4), 1095-1115.
- [96] Rodrigues, R. (2020). Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, 4, 100005.
- [97] Sarker, I. H. (2022). Ai-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems. *SN Computer Science*, 3(2), 158.
- [98] Sawhney, N., & Torres, A. P. G. (2022). Devising Regulatory Sandboxes and Responsible Practices for Designing AI-based Services in the Finnish Public Sector. on AI Compliance Mechanism (WAICOM 2022), 8.
- [99] Scatiggio, V. (2020). Tackling the issue of bias in artificial intelligence to design ai-driven fair and inclusive service systems. How human biases are breaching into ai algorithms, with severe impacts on individuals and societies, and what designers can do to face this phenomenon and change for the better.
- [100] Schiff, D. S., Schiff, K. J., & Pierson, P. (2022). Assessing public value failure in government adoption of artificial intelligence. *Public Administration*, 100(3), 653-673.
- [101] Schmidt, E., Work, B., Catz, S., Chien, S., Darby, C., Ford, K., ... & Moore, A. (2021). National security commission on artificial intelligence (ai). National Security Commission on Artificial Intelligence.
- [102] Sen, P. L. (2023). The Role of Government Policies and Initiatives in Driving IoT Adoption among Creative Industries Stakeholders in China. *Journal of Digitainability, Realism & Mastery (DREAM)*, 2(03), 67-76.
- [103] Sharma, A., Gupta, S., & Dubey, S. K. (2024, March). Analysis on Symptoms Driven Disease Risk Assessment using Artificial Intelligence Approach. In 2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 1-7). IEEE.
- [104] Shava, E., & Vyas-Doorgapersad, S. (2022). Fostering digital innovations to accelerate service delivery in South African Local Government. *International Journal of Research in Business and Social Science (2147-4478)*, 11(2), 83-91.
- [105] Shin, D., & Park, Y. J. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, 98, 277-284.
- [106] Song, M., Xing, X., Duan, Y., Cohen, J., & Mou, J. (2022). Will artificial intelligence replace human customer service? The impact of communication quality and privacy risks on adoption intention. *Journal of Retailing and Consumer Services*, 66, Article 102900. <https://doi.org/10.1016/j.jretconser.2021.102900>
- [107] Spatola, N. (2024). The efficiency-accountability tradeoff in AI integration: Effects on human performance and over-reliance. *Computers in Human Behavior: Artificial Humans*, 2(2), 100099.
- [108] Srinivasan, K. R., Abd Rahman, N. H., & Ravana, S. D. (2025). AI-Driven Decision Making Enhancing Public Sector Efficiency and Effectiveness. In *Harnessing AI, Blockchain, and Cloud Computing for Enhanced e-Government Services* (pp. 63-92). IGI Global Scientific Publishing.
- [109] Stahl, B. C., & Stahl, B. C. (2021). Ethical issues of AI. *Artificial Intelligence for a better future: An ecosystem perspective on the ethics of AI and emerging digital technologies*, 35-53.
- [110] Stoica, I., Song, D., Popa, R. A., Patterson, D., Mahoney, M. W., Katz, R., ... & Abbeel, P. (2017). A berkeley view of systems challenges for ai. arXiv preprint arXiv:1712.05855.
- [111] Susar, D., & Aquaro, V. (2019, April). Artificial intelligence: Opportunities and challenges for the public sector. In *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance* (pp. 418-426).
- [112] Trajkovski, G. (2024). Bridging the public administration-AI divide: A skills perspective. *Public Administration and Development*, 44(5), 412-426.
- [113] Tsuruta, S., Wakimoto, K., Nakamura, T., Siahpour, S., Miller, M., Taco, J., & Lee, J. (2023, September). Advancing predictive maintenance: A study of domain adaptation for fault identification in gearbox components. In *PHM Society Asia-Pacific Conference (Vol. 4, No. 1)*.
- [114] Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4), 101509.
- [115] Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4), 101509.
- [116] Van Noordt, C., & Misuraca, G. (2022). Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union. *Government Information Quarterly*, 39(3), 101714.
- [117] Verma, S., & Singh, V. (2022). Impact of artificial intelligence-enabled job characteristics and perceived substitution crisis on innovative work behavior of employees from high-tech firms. *Computers in Human Behavior*, 131(8), Article 107215. <https://doi.org/10.1016/j.chb.2022.107215>

- [118] Wang, K. H., & Lu, W. C. (2025). AI-induced job impact: Complementary or substitution? Empirical insights and sustainable technology considerations. *Sustainable Technology and Entrepreneurship*, 4(1), 100085.
- [119] Wang, X., & Wu, Y. C. (2024). Balancing innovation and Regulation in the age of generative artificial intelligence. *Journal of Information Policy*, 14.
- [120] Wilson, C., & Van Der Velden, M. (2022). Sustainable AI: An integrated model to guide public sector decision-making. *Technology in Society*, 68, 101926.
- [121] Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector—applications and challenges. *International Journal of Public Administration*, 42(7), 596-615.
- [122] Wirtz, B.W., Weyerer, J.C. and Geyer, C. (2018) ‘Artificial Intelligence and the Public Sector—Applications and Challenges’, *International Journal of Public Administration*, 42(7), pp. 596–615. doi:10.1080/01900692.2018.1498103.
- [123] Yamin, M. M., Ullah, M., Ullah, H., & Katt, B. (2021). Weaponized AI for cyber attacks. *Journal of Information Security and Applications*, 57, 102722.
- [124] Yang, L. W. Y., Ng, W. Y., Lei, X., Tan, S. C. Y., Wang, Z., Yan, M., ... & Ting, D. S. W. (2023). Development and testing of a multi-lingual Natural Language Processing-based deep learning system in 10 languages for COVID-19 pandemic crisis: A multi-center study. *Frontiers in Public Health*, 11, 1063466.
- [125] Zafar, A., & Villeneuve, S. (2018). Adopting AI in the Public Sector: Turning risks into opportunities through thoughtful design. Brookfield Institute for Innovation+ Entrepreneurship (Blog). Available at: <https://brookfieldinstitute.ca/commentary/adopting-ai-in-the-public-sector> April, 25, 2018.
- [126] Zorrilla, M., & Yebenes, J. (2022). A reference framework for the implementation of data governance systems for =industry 4.0. *Computer Standards & Interfaces*, 81, 103595.
- [127] Zuiderwijk, A., Chen, Y. C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, 38(3), 101577.
- [128] Zuiderwijk, A., Chen, Y. C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, 38(3), 101577.