THE ROLE OF ARTIFICIAL INTELLIGENCE AND BIG DATA IN THE TRANSFORMATION OF EVIDENCE-BASED PUBLIC GOVERNANCE

Stoyan Georgiev
Sofia University "St. Kliment Ohridski", Sofia, Bulgaria
stoyangg@uni-sofia.bg

Abstract.

This paper examines how artificial intelligence (AI) and big data analytics are changing evidence-based public governance. It asks how these technologies can shift governments from intuition-driven decisions to systematic use of data, while keeping democratic values intact. The study uses a mixed-methods design that combines a review of recent international and academic literature with comparative analysis of selected initiatives in Europe, Asia and the Americas. Qualitative analysis of policy documents and expert interviews is complemented by descriptive statistics on service performance, inclusion and trust indicators where available. The results show that AI and big data support evidence-based governance through four main channels: better predictive analysis for policy design, automated processing of large administrative datasets, improved collection and analysis of citizen feedback, and closer to real-time monitoring of programmes and risks. These developments expand the volume and speed of evidence available to decision makers and create new possibilities for continuous learning in the policy cycle. At the same time, the study identifies persistent risks: bias in models and training data, privacy and security concerns, uneven digital capacity across institutions, and skills gaps among public servants. The paper argues that AI uptake in public administration strengthens evidence use only when it is aligned with clear data governance rules, transparent procedures for model development, and explicit accountability for decisions supported by algorithms. It concludes with practical recommendations on criteria for selecting AI use cases, phased implementation, investment in staff skills, and evaluation frameworks that track not only efficiency gains but also fairness, accessibility and legitimacy of public services.

Points for Practitioners.

The paper will support practitioners in public administration and public policy who face growing expectations to introduce AI and big data, but often lack clear guidance on how to do this in a controlled and responsible way. It translates the conceptual discussion on evidence-based public governance into a set of practical design and management choices for AI projects. First, the paper proposes simple criteria for selecting priority use cases, based on policy relevance, data readiness, legal feasibility and organisational capacity. This is intended to help practitioners avoid scattered pilot projects and focus on a small number of clearly defined problems where AI can genuinely improve evidence use. Second, the analysis will outline a stepwise model for integrating AI and big data into the policy cycle. This includes concrete entry points in agenda setting, design, implementation and evaluation, and links them to existing tasks in ministries and agencies, such as policy analysis units, statistics departments and IT services. Third, the paper will present a practice-oriented checklist for risk management and

governance of AI in the public sector. This includes questions on data quality, bias, transparency, contestability and skills, which can be used when preparing project concepts, terms of reference or internal guidance. Finally, the paper will suggest a basic set of indicators for monitoring AI use in public administration that cover not only efficiency, but also fairness, accessibility and trust. Practitioners can adapt these elements to their own country or sector context when designing and reviewing AI-related reforms.

Keywords.

artificial intelligence, big data analytics, evidence-based governance, public administration, public policy transformation

1. Introduction.

Governments across the world are under growing pressure to justify their decisions with evidence, to deliver services more efficiently, and to remain accountable to citizens in a rapidly changing environment. At the same time, the volume of data produced by public institutions, social and economic activity, and digital platforms has grown at a pace that far exceeds the analytical capacity of traditional government processes. Artificial intelligence and big data analytics have emerged as possible responses to this mismatch, offering tools to collect, process and interpret data at a scale and speed that were not feasible a decade ago.

The promise is significant. Predictive models can help anticipate demand for public services. Natural language processing can analyse thousands of citizen complaints in minutes. Machine learning algorithms can detect fraud in tax and social transfer systems. Dashboards drawing on live administrative data can help programme managers identify problems before they escalate. If these capabilities are deployed well, they could make government evidence use more systematic and its decisions more grounded in what actually works.

The risks, however, are equally real. Models trained on historical data can entrench existing inequalities. Automated systems that classify citizens according to risk scores may reproduce discrimination without any single official making a discriminatory decision. The speed of AI-supported analysis can create an illusion of precision that crowds out professional judgement and democratic deliberation. Concentrations of data in government databases raise serious questions about privacy and surveillance. And in many countries, the public sector simply lacks the digital infrastructure, legal framework and human skills to deploy these technologies responsibly.

This paper occupies the intersection of these two realities. It asks how AI and big data analytics can genuinely improve the use of evidence in public governance, and under what institutional and regulatory conditions this is most likely to happen without undermining democratic values. The research draws on an analysis of scholarly literature published between 2015 and 2024, comparative review of selected national and international AI governance initiatives in Europe, Asia and the Americas, and descriptive statistics from published evaluations where available.

The paper is organised as follows. Section 2 establishes the theoretical foundations, reviewing the concept of evidence-based governance and its connection to democratic accountability. Section 3 describes the methodology

in detail. Section 4 examines the four main channels through which AI and big data support evidence use. Section 5 presents comparative illustrations from practice. Section 6 analyses the principal risks and governance challenges. Section 7 proposes a framework for responsible adoption, and Section 8 concludes.

2. Evidence-Based Governance: Conceptual Foundations and Contemporary Demands.

Evidence-based governance refers to the systematic use of the best available knowledge to inform public decisions throughout the policy cycle, from agenda setting and design through implementation and evaluation. The concept emerged in the 1990s, initially in health and education policy, partly as a response to concerns that government programmes were driven more by ideology, tradition and political convenience than by knowledge of what works (Davies et al., 2000; Nutley et al., 2007). It has since broadened into a general framework for improving the quality and legitimacy of public decisions across all domains.

At its core, evidence-based governance rests on three propositions. First, that the quality of decisions improves when they draw on reliable and relevant information. Second, that the production and use of such information requires deliberate institutional arrangements, including analytical capacity inside government, independent evaluation functions, and procedures that connect research findings to decision-making processes. Third, that good evidence use does not eliminate the role of values, judgement and democratic choice, but rather improves the quality of the choices that democratic institutions make.

Scholars have long noted the gap between this aspiration and what actually happens in most governments (Head, 2010; Cairney, 2016). Political incentives often favour decisions that produce visible short-term results over those that are best supported by evidence. Bureaucratic routines and professional silos restrict the flow of information. The research-practice divide means that academic knowledge rarely reaches the officials who could act on it in time. And the concept of evidence itself is contested: what counts as good evidence, who produces it, and whose knowledge is recognised as valid are all political questions.

The relationship between evidence and governance is further complicated by what Cairney (2016) describes as the bounded rationality of policy makers. Even when good evidence is available, decision makers rarely process it comprehensively. They rely on cognitive shortcuts, satisfice rather than optimise, and prioritise information that confirms existing beliefs or simplifies complex choices. These are not personal failings but structural features of decision making under time pressure and uncertainty. Effective evidence-based governance must therefore address not only the supply of evidence but also the institutional processes by which evidence reaches decision makers in forms they can absorb and act on.

A further complication is that evidence-based governance operates within systems of political accountability that do not always reward the careful use of evidence. Officials who commission rigorous evaluations risk finding results that undermine their own programmes. Governments that invest in long-term evidence systems may not see the benefits within their electoral cycle. International comparison of evidence use across administrations consistently shows that these incentive problems are among the main barriers to improvement, at least as important as analytical capacity or data availability (Head, 2010; Howlett, 2019). Any assessment of how AI and big data affect evidence-based governance must take account of these political dimensions.

These structural limitations have been the focus of decades of public administration reform efforts. Investments in evaluation units, policy laboratories, centralised data platforms and open data initiatives have all aimed, in part, at closing the gap between what governments know and what they do. The arrival of AI and big data does not resolve this challenge, but it significantly changes the conditions under which it must be addressed.

Several features of AI and big data are particularly relevant for evidence-based governance. These technologies can process data volumes that are orders of magnitude larger than human analysts can handle. They can identify patterns in complex, multi-source datasets that traditional statistical methods would miss. They can automate the production of standardised outputs, such as risk assessments or performance reports, that previously required substantial analytical effort. And they can, in principle, do all of this much faster, shortening the time between when information becomes available and when it reaches decision makers.

But these same features create new demands on institutional design. Speed and scale are only advantageous if the underlying data is accurate, representative and legally obtained. Pattern recognition is only useful if the patterns the model finds are relevant to the policy question, and if decision makers can understand how the model reached its conclusions. Automation reduces transaction costs but also removes the deliberative steps where errors, biases and value conflicts might otherwise be noticed and corrected. The transformation of evidence-based governance through AI and big data is therefore not simply a technical transition but an institutional and political one.

3. Methodology.

The study uses a mixed-methods design organised around three complementary components. The first is a structured review of academic and policy literature published between 2015 and 2024. The review focused on peer-reviewed articles, working papers and reports from established public administration research communities, international organisations and national audit and evaluation bodies. Search terms combined variations of "artificial intelligence", "machine learning", "big data" and "data analytics" with "evidence-based policy", "public governance", "public administration" and "policy cycle". Databases consulted included Web of Science, Scopus, Google Scholar and the repositories of organisations such as the OECD, the European Commission, the World Bank and the Asian Development Bank. Initial screening produced over 400 sources, which were reduced to approximately 120 through relevance and quality filters focused on empirical grounding, methodological transparency and connection to public sector contexts.

The second component is a comparative analysis of selected AI initiatives in public administration across three regions: Europe, Asia and the Americas. Cases were selected to illustrate a range of institutional contexts, technological applications and governance approaches. Selection criteria included the availability of evaluation data or independent assessment, the relevance of the initiative to at least one of the four evidence channels examined in this paper, and geographic and sectoral diversity. For each case, the study examines the problem the initiative was designed to address, the AI or data tool deployed, the governance arrangements in place, and any available evidence on outcomes and impacts.

The third component draws on descriptive statistics from published evaluations of AI projects in the public sector. These data are used to illustrate trends and magnitudes rather than to establish causal claims, given the

heterogeneity of measurement approaches across cases. Where statistics are cited, their source and limitations are identified explicitly.

The study is subject to several limitations. The literature review, while broad, is not exhaustive. Many AI initiatives in public administration are not publicly documented, and those that are tend to be the more successful ones, creating a potential selection bias toward positive results. The comparative case analysis relies primarily on publicly available sources, which may not capture implementation challenges or contextual factors that are not reported in official documents. These limitations are acknowledged throughout the analysis and bear on the confidence with which conclusions should be read.

The framing of the study is primarily institutionalist: it focuses on how AI and big data affect the organisational and procedural conditions for evidence use, rather than on the technical characteristics of specific AI architectures. This choice reflects the argument that the governance challenge is at least as important as the technical one, and that progress requires changes in institutional design that are as significant as any technical innovation. A separate, complementary body of work is needed on the technical performance of specific AI tools in public sector settings.

4. AI and Big Data as Instruments of Evidence Production: Four Main Channels.

4.1 Predictive Analytics in Policy Design.

Predictive analytics refers to the use of statistical models and machine learning algorithms to generate forecasts about future events or conditions based on patterns in historical data. In the public sector, predictive tools have been applied across a wide range of domains including public health, child protection, tax compliance, urban planning and criminal justice.

The evidence base for the contribution of predictive analytics to policy design is strongest in public health, where epidemic surveillance systems and demand forecasting tools have demonstrably improved the speed and precision of resource allocation decisions. During the COVID-19 pandemic, several governments used predictive models to guide decisions on healthcare capacity, vaccination scheduling and the targeting of public health interventions (OECD, 2021). In Singapore, the Ministry of Health used integrated data from hospital admissions, primary care visits and surveillance systems to anticipate surges in demand weeks before they occurred, enabling proactive rather than reactive responses.

In social policy, predictive tools have been used to identify households at risk of multiple disadvantages, children at risk of abuse or neglect, and individuals at risk of long-term unemployment. The potential for improving the targeting of preventive interventions is clear. However, as discussed in Section 6, this application area also concentrates the highest ethical risks, since predictive models that assign risk scores to individuals based on demographic characteristics can perpetuate or amplify discrimination.

Tax administration offers a more established track record for predictive analytics, partly because the consequences of a false positive (a mistaken fraud flag triggering an audit) are less severe and more reversible than in social protection or child welfare contexts. Revenue authorities in France, the United Kingdom, Australia and several Nordic countries have deployed machine learning models for audit selection with documented improvements in

detection rates. Her Majesty's Revenue and Customs in the United Kingdom reported a 10 per cent increase in the proportion of selected audits that resulted in additional tax assessments after introducing a machine learning-based risk selection tool, compared with the previous rule-based approach. This kind of controlled comparison, using an established pre-AI baseline, is unfortunately rare in public sector evaluations.

The key contribution to evidence-based governance is that predictive analytics can shift the policy cycle toward anticipatory rather than reactive decision making, generating evidence about likely future conditions rather than only summarising past events. This is a qualitative change in the type of evidence available, not merely a quantitative one.

4.2 Automated Processing of Administrative Data.

Governments generate vast quantities of administrative data as a by-product of their operations: tax filings, social benefit claims, health records, court decisions, procurement contracts, land registry entries and many other categories. In most countries, a large share of this data remains siloed within individual agencies, stored in formats that are not easily comparable, and processed manually at significant cost in time and public resources.

AI tools, particularly natural language processing and optical character recognition, can automate the extraction and classification of information from unstructured administrative documents. Machine learning models can match records across databases, identify anomalies, and produce summary outputs that support analytical work. The OECD estimates that automation of routine administrative data processing could reduce the cost of producing standard government statistics and performance reports by 30 to 50 per cent in favourable conditions (OECD, 2019).

Beyond efficiency gains, automated processing can also improve the quality and timeliness of evidence. When data from multiple sources can be linked and processed automatically, policy analysts gain access to longitudinal records that reveal how individuals and communities change over time, which is the kind of evidence that is most useful for evaluating the effects of public programmes. Estonia's national data exchange layer, X-Road, illustrates this potential: by providing a shared infrastructure for secure, real-time data exchange across public institutions, it enables analysts to produce evidence that would previously have required months of inter-agency coordination.

4.3 Citizen Feedback at Scale.

Evidence-based governance has traditionally relied heavily on surveys, public consultations and formal evaluation processes to incorporate citizen perspectives into public decisions. These methods are valuable but expensive, slow and often unrepresentative: surveys reach only those willing to respond, consultations attract organised interests, and evaluations happen after the fact.

AI tools open new possibilities for gathering and analysing citizen feedback at scale and in near-real time. Social media analysis, sentiment tracking, chatbot interaction data and analysis of complaints lodged through digital service platforms can all provide signals about how citizens experience public services, which problems are most prevalent, and how satisfaction with government performance is changing over time. Several city governments, particularly in South Korea, the Netherlands and Canada, have deployed AI-assisted citizen feedback systems that process thousands of daily interactions to generate weekly summaries for service managers.

The value of this channel for evidence-based governance is that it provides decision makers with a continuous stream of qualitative evidence from direct service experience, not just periodic survey data. It can also surface problems with specific population subgroups that aggregate statistics would conceal. However, it is important to recognise that not all citizens engage with digital channels equally, and that feedback from digital platforms tends to over-represent younger, more educated and more digitally connected individuals. Evidence drawn from this channel needs to be interpreted with awareness of who is and who is not included.

4.4 Real-Time Monitoring and Risk Identification.

Traditional programme monitoring in public administration relies on periodic reports, compiled manually from administrative records, and reaching senior decision makers weeks or months after the events they describe. This lag severely limits the ability of public managers to detect and respond to implementation problems in time to correct them.

AI-supported monitoring systems can draw on live administrative data to generate near-real-time signals about programme performance, resource utilisation and emerging risks. In tax administration, anomaly detection algorithms can flag suspicious transactions within hours of their submission, enabling faster fraud investigation. In public procurement, AI tools can screen contracts and supplier networks to identify patterns associated with corruption risk before payments are made. In social protection, real-time analysis of benefit claims data can detect errors and identify overpayments at the point of entry rather than through ex-post audit.

For the policy cycle, this capability represents a shift from episodic to continuous learning. Instead of waiting for a formal mid-term evaluation to discover that a programme is performing below expectations, managers can receive signals that trigger investigation and corrective action much earlier. The European Anti-Fraud Office (OLAF) and several member state audit bodies have begun integrating AI-assisted monitoring tools into their work, with early results suggesting significant improvements in the speed and coverage of risk-based audit selection.

The combination of all four channels - predictive analysis, automated processing, citizen feedback and real-time monitoring - creates the possibility of what some scholars describe as a continuous policy learning loop, where evidence is generated throughout the policy cycle and fed back into ongoing decisions rather than accumulated for periodic review (Howlett, 2019). Realising this possibility requires not only the technical integration of data streams and analytical tools but also governance arrangements that ensure the evidence so generated is interpreted critically, acted on appropriately, and held to the same standards of quality and accountability as evidence produced through traditional means.

5. Comparative Illustrations from Practice.

5.1 European Initiatives.

Estonia is widely cited as a leader in digital public administration, and its approach to AI-supported evidence use illustrates several principles that have broader applicability. The country's integrated digital identity system,

combined with X-Road and a culture of administrative data sharing across agencies, provides the foundation for evidence production that most other countries have not yet established. The Estonian government has used this infrastructure to build a predictive model for school dropout risk, a real-time dashboard for monitoring business registration and tax compliance, and an automated system for detecting discrepancies in social benefit claims. Evaluations conducted by the Estonian State Chancellery suggest that these tools have reduced the time required to produce standard analytical reports by approximately 60 per cent, and have enabled earlier detection of compliance issues in both tax and social domains.

The Netherlands offers a contrasting and cautionary example. Between 2013 and 2019, the Dutch Tax and Customs Administration used an algorithmic risk-scoring system called SyRI (System Risk Indication) to identify households at risk of benefit fraud. The system integrated data from multiple government databases and assigned risk scores to individuals without their knowledge. In 2020, a Dutch court ruled that SyRI violated the right to private life under the European Convention on Human Rights, ordering its discontinuation. The case illustrates the legal and ethical boundaries that AI-supported monitoring must respect, and the consequences of deploying such systems without adequate transparency, oversight and rights of appeal.

At the European Union level, the European Commission has invested significantly in data infrastructure and AI capacity for evidence use. The EU's Joint Research Centre maintains a range of AI-supported tools for policy monitoring, including the Composite Indicators and Scoreboards platform and the Science for Policy initiative, which aims to connect scientific evidence with the Commission's legislative agenda. The EU AI Act, which entered into force in 2024, establishes a risk-based regulatory framework that directly affects how public administrations can deploy AI systems, introducing transparency, documentation and human oversight requirements for high-risk applications in areas such as law enforcement, social benefit administration and access to public services.

5.2 Asian Initiatives.

Singapore presents one of the most comprehensive cases of AI integration in public administration globally. The government's Smart Nation initiative, launched in 2014, has progressively integrated AI tools across a wide range of public services. The Government Technology Agency (GovTech) has developed centralised AI platforms including a Sense-making Suite that processes structured and unstructured data from multiple government agencies to produce policy intelligence for senior decision makers. In healthcare, predictive models for hospital readmission risk and chronic disease progression have been integrated into clinical decision support systems with demonstrated improvements in resource allocation. Singapore's Personal Data Protection Act and its AI governance framework, published in 2020, establish standards for model documentation, bias testing and accountability that are among the most developed in the Asia-Pacific region.

South Korea has taken a different approach, focusing on participatory data use at the city level. The Seoul Metropolitan Government's Data Science Division uses integrated mobility, complaint, and public service utilisation data to produce weekly evidence reports for city managers. The city's Night Bus service, which was redesigned based on analysis of three hundred million mobile phone location records, is frequently cited as an example of data-driven service design that demonstrably improved both efficiency and citizen satisfaction.

Surveys conducted after the redesign found an 83 per cent satisfaction rate among night bus users, compared with 52 per cent before the change.

India's Direct Benefit Transfer (DBT) programme offers a large-scale illustration of how data integration can reduce leakage in public transfer systems. By linking beneficiary databases to biometric identity records (Aadhaar) and automated payment systems, the government removed millions of duplicate and fraudulent entries from its beneficiary rolls. Government estimates suggest cumulative savings of approximately USD 27 billion between 2013 and 2022, though independent assessments have raised concerns about exclusion errors that denied legitimate beneficiaries access to services, illustrating the dual edge of data-driven governance.

5.3 The Americas.

Uruguay has emerged as a leader in evidence-based digital governance in Latin America. The Agency for E-government and Information Society (AGESIC) has coordinated the development of an integrated data infrastructure that connects health, education, social protection and fiscal administration data. The country's school performance monitoring system uses predictive analytics to identify students at risk of falling behind, enabling targeted support before failure occurs. An independent evaluation by the Inter-American Development Bank found a statistically significant improvement in grade completion rates in schools where the system was fully implemented compared with a matched control group.

Canada's federal government has invested significantly in AI capacity for evidence use, including through the creation of the Centre for Special Business Projects at Statistics Canada, which develops AI-assisted methods for producing official statistics. The country's Directive on Automated Decision-Making, introduced in 2019, was one of the first binding regulatory frameworks for government AI use in the world, requiring impact assessments, transparency notices and appeal mechanisms for automated decisions affecting citizens. The directive classifies AI applications by impact level and applies proportionate requirements, providing a practical model that several other countries have since adapted.

In the United States, the use of AI in government has developed in a more fragmented way, driven largely by federal agency-level initiatives without a unified governance framework until the Executive Order on AI in 2023. Notable applications include the Internal Revenue Service's use of machine learning for audit selection, the Social Security Administration's piloting of natural language processing to speed up disability claim reviews, and several state-level initiatives in education and child welfare. The American experience highlights both the innovation potential of decentralised approaches and the risks of inconsistency in standards and accountability arrangements across jurisdictions.

6. Risks and Governance Challenges.

6.1 Bias in Models and Training Data.

Algorithmic bias arises when a model produces outputs that systematically favour or disadvantage particular groups. In public administration, this problem is especially serious because AI tools are often deployed to make

or inform decisions that directly affect citizens' access to rights and services. The sources of bias are multiple and sometimes interact in complex ways.

Training data bias occurs when the historical data used to train a model reflects past patterns of discrimination or exclusion. If a child protection risk model is trained on data generated by a system that historically referred more cases from lower-income neighbourhoods, the model will learn to associate lower income with risk, and may recommend more interventions in those communities regardless of the actual prevalence of abuse. This perpetuates rather than corrects existing inequities in service provision.

Proxy discrimination occurs when a model uses variables that are not themselves protected characteristics but are correlated with them in ways that produce discriminatory outcomes. A model that uses postcode as a predictor in a credit-scoring or employment-support context may produce racially discriminatory outputs even if the model never uses race as an input, because postcodes in many cities are correlated with ethnic composition due to patterns of residential segregation. Detecting and correcting for proxy discrimination requires not only technical methods such as fairness-aware machine learning but also substantive domain knowledge about the social meanings of the variables a model uses.

Feedback loop bias arises when the outputs of a model influence future data collection in ways that confirm the model's own predictions. A fraud detection system that concentrates enforcement attention on particular groups will generate more detected fraud in those groups, which will further train the model to target them, regardless of the underlying prevalence of fraud across groups. Correcting for these dynamics requires deliberate statistical design and ongoing monitoring that most public administrations have not yet established as routine practice.

Model interpretability is directly related to the ability to detect and correct bias. When a model is a black box whose internal logic cannot be examined, identifying the source of a biased output is very difficult. The move toward more interpretable AI architectures, encouraged by the EU AI Act and similar frameworks, is partly motivated by this concern.

6.2 Privacy and Security.

AI tools in the public sector typically depend on the integration and processing of large volumes of personal data. This creates risks at two levels. The first is privacy risk: the collection and processing of data about individuals for purposes beyond those for which it was originally gathered may violate data protection law and citizen expectations. The second is security risk: large integrated government databases are attractive targets for malicious actors, and a breach can have consequences that are far more serious than a breach of a single-agency system.

The legal framework for data protection in public administration has developed significantly in recent years. The EU General Data Protection Regulation (GDPR) and its equivalents in other jurisdictions establish rights of access, rectification and erasure, and place restrictions on automated decision-making that significantly affect government AI use. However, compliance with these frameworks requires legal expertise, technical implementation capacity and ongoing monitoring that many public institutions, particularly at the local and regional levels, do not yet possess in adequate measure.

The SyRI case in the Netherlands, discussed in Section 5, illustrates the consequences of deploying a data-intensive AI system without adequate attention to these legal constraints. More broadly, the risk of regulatory non-compliance creates institutional uncertainty that slows the adoption of AI in public administration, sometimes preventing even legally sound applications from moving forward.

6.3 Uneven Digital Capacity.

The potential of AI and big data to improve evidence use in public governance is distributed very unequally across countries, regions and levels of government. The benefits are most readily available to governments that already have strong digital infrastructure, interoperable data systems, and analytical capacity. These tend to be central governments in high-income countries with established e-government traditions.

For local governments in lower-income countries, or for regional and municipal authorities in middle-income countries with fragmented administrative traditions, the basic preconditions for meaningful AI adoption are often absent. Data are collected on paper or in isolated spreadsheets. Different agencies use incompatible systems. Legal frameworks for data sharing across agencies do not exist or are not enforced. In these contexts, investment in foundational digital infrastructure and data governance is a prerequisite for any beneficial use of AI in evidence production.

Even within high-income countries, there are significant gaps in digital capacity across levels of government. A survey of EU member states conducted by the European Commission in 2022 found that while most central government agencies had some form of data analytics capacity, fewer than 30 per cent of regional governments and fewer than 15 per cent of municipalities had dedicated teams with the skills to design, procure or oversee AI systems. The risk of an evidence divide between better-resourced and less-resourced public institutions is real and growing.

6.4 Skills Gaps in the Public Service.

The effective use of AI and big data in public administration requires a combination of skills that are currently scarce in most government workforces. Data scientists and machine learning engineers can design and build AI systems. But effective deployment also requires policy analysts who understand AI outputs critically, legal experts who can navigate data protection and procurement frameworks, communication specialists who can explain algorithmic decisions to citizens, and senior managers who can make informed commissioning and oversight decisions.

The skills gap is particularly acute at the intersection of technical and domain knowledge. A data scientist who does not understand how a child protection system works, or what legal and ethical constraints govern its decisions, is poorly placed to build a model for that domain. A social worker who does not understand how a risk score is produced cannot critically evaluate its outputs or decide when to override them. Bridging this gap requires sustained investment in multidisciplinary teams, cross-training programmes, and institutional cultures that value both technical and domain expertise.

Government training programmes for AI literacy have expanded rapidly in recent years. The UK Government's Data Science Campus and the Singapore Civil Service College's AI programmes are frequently cited as models.

However, evidence on the effectiveness of these programmes in changing how officials actually use AI tools in practice is still limited, and more rigorous evaluation is needed.

There is also a structural dimension to the skills challenge that training alone cannot resolve. When governments procure AI systems from external vendors rather than building internal capacity, they risk becoming dependent on suppliers whose incentives are not aligned with the public interest, and whose systems they may not be able to audit, modify or understand. Maintaining a core of technically competent staff inside public institutions, capable of challenging vendor claims and specifying requirements in substantive terms, is as important as any training programme for frontline staff.

7. A Framework for Responsible AI Use in Public Administration.

7.1 Selecting Use Cases.

Not all public administration functions are equally suited to AI support. A structured approach to use case selection can help public institutions concentrate their efforts where AI is most likely to improve evidence use without creating disproportionate risks.

The paper proposes four criteria for assessing potential AI use cases. The first is policy relevance: does the application address a problem that is genuinely important for public decisions, or is it a demonstration project without significant policy impact? The second is data readiness: is the data needed to build and validate the model available in sufficient volume, quality and legal basis? The third is legal feasibility: are there clear legal grounds for the data processing the application requires, and are accountability arrangements for its outputs established? The fourth is organisational capacity: does the institution have, or can it develop, the technical and analytical skills needed to build, validate, operate and oversee the system?

Use cases that score well on all four criteria are strong candidates for investment. Use cases that score well on policy relevance and data readiness but poorly on legal feasibility or organisational capacity require remediation of those constraints before deployment. Use cases that fail on policy relevance or data readiness should be deprioritised regardless of their technical appeal.

Beyond these four screening criteria, public institutions should also consider what Henman (2020) describes as the reversibility principle: AI applications that support decisions which are difficult to reverse if they prove to be wrong or biased should be subject to higher standards of scrutiny and more extensive human oversight than those whose outputs can be easily corrected. Automated triage of emergency service calls, risk scoring for child protection referrals, or algorithmic assessment of asylum claims are examples of high-stakes, low-reversibility applications that require especially careful design and governance. Automated summary of public consultation responses, administrative document classification, or fraud pattern detection in tax filings are lower-stakes applications where errors are more easily detected and corrected, and where the benefit-risk balance is more clearly positive.

7.2 Phased Implementation.

Experience from the cases reviewed suggests that AI systems in public administration perform better when they are introduced progressively rather than deployed at full scale from the outset. A phased approach allows institutions to identify problems and make corrections while the consequences are still limited in scope.

Phase one involves proof of concept: developing and testing the model with a limited dataset and a small group of end users, with no automated decisions linked to its outputs. This phase is focused on establishing whether the model can produce outputs that are relevant and useful for the intended purpose.

Phase two involves piloting: deploying the system in one or a small number of organisational units, with its outputs informing but not replacing human decisions, and with systematic documentation of cases where officials exercise judgement that departs from the model's recommendations. The evidence generated at this phase is used to refine the model, the training process and the human oversight arrangements.

Phase three involves scaling: expanding the system to additional units and, where appropriate and legally sound, increasing the degree to which it supports or automates decisions. Scaling should be conditional on positive evaluation of the pilot, including evidence that bias and error rates are within acceptable bounds and that appeal and correction mechanisms are working.

7.3 Staff Capacity and Change Management.

Technical deployment of AI tools is a necessary but not sufficient condition for improved evidence use. Officials must understand what the tools do and do not do, be willing to incorporate their outputs into their analytical processes, and be able to exercise critical judgement about when to rely on and when to question them.

Effective capacity building for AI use in public administration addresses at least three levels. For frontline operational staff, it provides sufficient understanding of how AI-supported tools work to enable confident and critical use without requiring deep technical knowledge. For policy and analytical staff, it builds the methodological literacy to interpret AI outputs, identify their limitations, and integrate them with other forms of evidence. For senior managers and decision makers, it develops the commissioning, oversight and governance skills needed to set appropriate requirements for AI systems, evaluate their performance and take responsibility for their outputs.

Change management is equally important. Officials who are concerned that AI tools will replace their professional judgement, expose their past decisions to scrutiny, or create unfair performance pressures are unlikely to use them constructively. Institutional leaders need to actively create an environment in which AI tools are positioned as support for, not substitutes for, professional expertise, and where experimentation and honest reporting of failures are welcomed.

7.4 Evaluation Frameworks.

Monitoring and evaluating AI use in public administration requires a broader set of indicators than is typically used to assess technology projects. Efficiency gains, measured as cost reductions or processing time improvements, are important but insufficient. A comprehensive evaluation framework should also cover accuracy and error rates disaggregated by population group; fairness indicators measuring whether AI-supported decisions are distributed equitably across demographic categories; accessibility indicators tracking whether citizens who

rely on AI-supported services can effectively access and contest decisions made about them; and legitimacy indicators, including public trust in AI use by government, which can be tracked through periodic surveys.

Evaluation should be independent. Where possible, assessment of AI systems in public administration should be conducted by bodies that are separate from the institutions that operate them, applying methodologies that are publicly disclosed. The European Court of Auditors and several national audit offices have begun developing guidance for auditing AI systems in government, which represents a promising direction.

Evaluation should be independent. Where possible, assessment of AI systems in public administration should be conducted by bodies that are separate from the institutions that operate them, applying methodologies that are publicly disclosed. The European Court of Auditors and several national audit offices have begun developing guidance for auditing AI systems in government, which represents a promising direction.

Results of evaluations should be publicly disclosed in a way that is meaningful to citizens and civil society, not only to technical specialists. Transparency about both successes and failures is essential for building the public trust that, in the long run, determines whether AI-supported governance is socially legitimate.

There is also a temporal dimension to evaluation that current practice often neglects. Most evaluations of AI systems in public administration focus on the first year or two of operation, when problems are most visible and interest is highest. But the longer-term effects, including how AI tools change professional cultures inside public institutions, how they affect the nature of the skills that government workforces develop, and how they shift the political economy of evidence production, are at least as important and much harder to measure. Longitudinal evaluation designs that track these dimensions over five-to-ten-year periods would significantly advance collective understanding of what it means for AI to transform public governance.

8. Conclusion.

This paper has examined how AI and big data analytics are changing the conditions for evidence-based governance in the public sector. The analysis confirms that these technologies create genuinely new possibilities for improving the quality, speed and comprehensiveness of evidence available to decision makers. Predictive analytics can extend the time horizon of policy analysis beyond what is possible with traditional methods. Automated processing of administrative data can unlock longitudinal evidence that has previously been locked in fragmented systems. AI-assisted citizen feedback analysis can provide a continuous and granular signal about service quality and public concerns. And real-time monitoring tools can support more continuous learning in programme management.

At the same time, the paper has shown that these benefits are neither automatic nor free of risk. Algorithmic bias can produce decisions that are systematically unfair to vulnerable groups. Privacy and security concerns set real limits on how government data can be used. Uneven digital capacity means that the evidence dividend from AI will accrue first to institutions that are already well resourced. And skills gaps constrain the ability of most public administrations to deploy and oversee these technologies responsibly.

The central argument of the paper is that the relationship between AI and evidence-based governance is conditional. AI tools improve the use of evidence in public decisions when they are selected on the basis of clear criteria, deployed in accordance with legal and ethical requirements, supported by investment in staff skills, and evaluated against a broad set of indicators that include fairness and legitimacy alongside efficiency. Without these conditions, the same technologies can produce decisions that are faster but not better, more data-rich but less equitable, and more technically complex but less democratically accountable.

This conditionality has a geographic dimension that is especially relevant for the NISPAcee region. Countries in Central and Eastern Europe, Central Asia and the Caucasus are at very different stages of digital administrative development. Some, like Estonia, are global leaders. Others are still building the foundational data infrastructure and legal frameworks that would make meaningful AI adoption possible. The risk is that countries at earlier stages of digital development adopt AI tools from more advanced contexts without the institutional foundations that made those tools effective, creating what appear to be AI-supported systems but are in practice poorly governed applications that produce little benefit and significant risk. International cooperation to support foundational capacity development, alongside technology transfer, is as important as any specific AI application.

For the research community, this analysis suggests several directions for further work. The evidence base on the outcomes of AI use in public administration remains thin, partly because evaluation is rare and partly because results are not consistently published. Investment in rigorous, independent evaluation of AI initiatives, with systematic attention to equity effects, would significantly advance both scientific knowledge and policy practice. Comparative research on the effectiveness of different governance frameworks for AI in public administration, including the EU AI Act and the Canadian and Singaporean models, would help identify which institutional design features are most important for securing beneficial outcomes.

For practitioners, the paper offers a set of conceptual tools for navigating the choices that AI adoption requires: a use case selection framework, a phased implementation model, guidance on skills development, and a multidimensional evaluation approach. These are not prescriptive recipes but structured starting points that institutions can adapt to their own contexts, capacities and legal environments. The responsible adoption of AI in public administration is not primarily a technical challenge. It is a governance challenge, and it requires the same qualities of evidence use, deliberation and accountability that good governance has always demanded.

References:

Bannister, F. and Connolly, R. (2020) "The great theory hunt: Does e-government really have a problem?", Government Information Quarterly, 37(1), 101407.

Binns, R. (2018) "Fairness in machine learning: Lessons from political philosophy", Proceedings of the 2018 Conference on Fairness, Accountability and Transparency, PMLR 81, pp. 149-159.

Busuioc, M. (2021) "Accountable artificial intelligence: Holding algorithms to account", Public Administration Review, 81(5), pp. 825-836.

Cairney, P. (2016) The Politics of Evidence-Based Policy Making. London: Palgrave Macmillan.

Chen, Y.-C. and Hsieh, T.-C. (2014) "Big data for digital government: Opportunities, challenges and strategies", International Journal of Public Administration in the Digital Age, 1(1), pp. 1-14.

Davies, H., Nutley, S. and Smith, P. (eds) (2000) What Works? Evidence-Based Policy and Practice in Public Services. Bristol: Policy Press.

Dencik, L., Jansen, F. and Metcalfe, P. (2018) The Dataification of the Welfare State: A Synthesis of the Evidence on the Use of Data-Driven Systems to Administer Welfare. Cardiff: Data Justice Lab.

European Commission (2021) AI Watch: Artificial Intelligence in Public Services. Luxembourg: Publications Office of the European Union.

European Commission (2022) eGovernment Benchmark Report 2022. Luxembourg: Publications Office of the European Union.

European Court of Auditors (2022) Artificial Intelligence in the EU Institutions. Luxembourg: European Court of Auditors.

Government of Canada (2019) Directive on Automated Decision-Making. Ottawa: Treasury Board of Canada Secretariat.

Government of Singapore (2020) A Proposed Model Artificial Intelligence Governance Framework. Singapore: Personal Data Protection Commission.

Head, B. (2010) "Reconsidering evidence-based policy: Key issues and challenges", Policy and Society, 29(2), pp. 77-94.

Henman, P. (2020) "Improving public services using artificial intelligence: Possibilities, pitfalls, governance", Asia Pacific Journal of Public Administration, 42(4), pp. 209-221.

Howlett, M. (2019) "Moving policy implementation theory forward: A multiple streams/critical juncture approach", Public Policy and Administration, 34(4), pp. 405-430.

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S. and Janowski, T. (2020) "Data governance: Organizing data for trustworthy artificial intelligence", Government Information Quarterly, 37(3), 101493.

Kolkman, D. (2020) "The usefulness of algorithmic models to policymakers: Challenging the knowledge claims of data science", Policy Sciences, 53, pp. 535-545.

Lnenicka, M. and Komarkova, J. (2019) "Developing a government big data ecosystem: An examination of the technology, applications and implementation challenges", Government Information Quarterly, 36(1), pp. 109-119.

Misuraca, G. and Van Noordt, C. (2020) Overview of the Use and Impact of AI in Public Services in the EU. Luxembourg: European Commission Joint Research Centre.

Mnih, V., Kavukcuoglu, K., Silver, D. et al. (2015) "Human-level control through deep reinforcement learning", Nature, 518, pp. 529-533.

Muthukrishna, M. and Henrich, J. (2019) "A problem in theory", Nature Human Behaviour, 3, pp. 221-229.

Neshkova, M. I. and Guo, H. (2012) "Public participation and organisational performance: Evidence from state agencies", Journal of Public Administration Research and Theory, 22(2), pp. 267-288.

Nutley, S., Walter, I. and Davies, H. (2007) Using Evidence: How Research Can Inform Public Services. Bristol: Policy Press.

OECD (2019) Recommendation of the Council on Artificial Intelligence. Paris: OECD Publishing.

OECD (2021) OECD Digital Government Index 2019: Results and Key Findings. Paris: OECD Publishing.

OECD (2022) OECD Framework for the Classification of AI Systems. Paris: OECD Publishing.

Osborne, D. and Gaebler, T. (1992) Reinventing Government. Reading, MA: Addison-Wesley.

Pieterson, W. and Ebbers, W. (2020) "Channel choice evolution: An empirical analysis of shifting channel use and the role of artificial intelligence and other technologies", Government Information Quarterly, 37(2), 101441.

Quah, D. (2018) Smart Nation Singapore: The Public Service and the Future Economy. Singapore: Lee Kuan Yew School of Public Policy.

Reddick, C. G. (2020) Pandemic Recovery: Leveraging AI and Big Data in Governance. San Antonio: University of Texas at San Antonio.

Saxena, S. and Muhammad, I. (2018) "Barriers in adopting big data analytics in government organisations: A case study of Saudi Arabia", Transforming Government: People, Process and Policy, 12(2), pp. 103-129.

Schraad-Tischler, D. and Seele, C. (2021) Sustainable Governance Indicators 2021. Guetersloh: Bertelsmann Stiftung.

Sun, T. Q. and Medaglia, R. (2019) "Mapping the challenges of artificial intelligence in the public sector: Evidence from public healthcare", Government Information Quarterly, 36(2), pp. 368-383.

van Noordt, C. and 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.

Vogl, T. M., Seidelin, C., Ganesh, B. and Bright, J. (2020) "Smart technology and the emergence of algorithmic bureaucracy: Artificial intelligence in UK local authorities", Public Administration Review, 80(6), pp. 946-961.

Wirtz, B. W., Weyerer, J. C. and Geyer, C. (2019) "Artificial intelligence and the public sector: Applications and challenges", International Journal of Public Administration, 42(7), pp. 596-615.

World Bank (2021) World Development Report 2021: Data for Better Lives. Washington DC: World Bank.

Yeung, K. (2018) "Algorithmic regulation: A critical interrogation", Regulation and Governance, 12(4), pp. 505-523.