Review paper / Обзорная статья
https://doi.org/10.51176/1997-9967-2025-2-85-103
МРНТИ 06.81.13
JEL: H83, L86, O33

Machine Learning in Public Governance: A Systematic Review of Applications, Trends and Challenges

Yeldar Nuruly[a,b]*, Galiya N. Sansyzbayeva[a], Laura Z. Ashirbekova[a], Samal K. Tazhiyeva[a]

[a] Al-Farabi Kazakh National University, 71 al-Farabi Ave., 050040, Almaty, Kazakhstan; [b] Centre for Sustainable Development in Central Asia, Al-Farabi Kazakh National University in partnership with the Hong Kong Polytechnic University, Almaty, Kazakhstan

For citation: Nuruly, Y., Sansyzbayeva, G. N., Ashirbekova, L.Z. & Tazhiyeva, S.K. (2025). Machine Learning in Public Governance: A Systematic Review of Applications, Trends and Challenges. Economy: strategy and practice, 20(2), 85-103, https://doi.org/10.51176/1997-9967-2025-2-85-103

ABSTRACT

Today, the active implementation of machine learning (hereinafter – ML) methods in public administration opens up new opportunities for forecasting, impact assessment and decision support, while simultaneously generating various challenges. The present study is aimed at a systematic review of scientific publications devoted to applying ML methods in the field of public administration, with an emphasis on identifying thematic areas, ethical and institutional challenges. The initial data set included 524 publications obtained using targeted search queries in the Scopus and Web of Science databases for the period 2014-2024. Data filtering was performed using SQLite, thematic mapping was performed in the VOSviewer environment, and metadata was structured using the Elicit tool and subsequent manual encoding. The analysis results allowed us to identify four functional areas of ML application in public administration: transparency and ethics, resource allocation and service provision, institutional design, and technical integration. Despite significant progress in the models' technical implementation and predictive accuracy, in many cases, mechanisms for equity, transparency, and citizen participation have been poorly implemented. The scientific novelty of the work lies in the interdisciplinary synthesis and development of a typology of institutional challenges that arise when implementing ML systems in public administration. The prospects for further research are related to the empirical validation of decisions, the development of ethical audit methods, and institutional training for responsible, sustainable, and contextually adaptive use of algorithmic tools in the public administration system.

KEYWORDS: Machine Learning, Public Administration, Public Policy, Technology Adoption, Strategic Planning, Digital Economy

CONFLICT OF INTEREST: the authors declare that there is no conflict of interest.

FINANCIAL SUPPORT: this research has been funded under the grant funded by the Committee of Science MSHE RK "Artificial Intelligence Technologies in Public Administration: Problems and Prospects for Application" (AP23487228).

Article history:
Received 25 March 2025
Accepted 29 May 2025
Published 30 June 2025

* **Corresponding author: Nuruly Y.** – PhD candidate, Senior Lecturer, Senior Research Fellow, Al-Farabi Kazakh National University, 71 al-Farabi Ave., 050040, Almaty, Kazakhstan, 87788797593, email: yeldar.nuruly@kaznu.edu.kz

Машинное обучение в государственном управлении: систематический обзор применений, трендов и вызовов

Нурулы Е.[a, b]*, Сансызбаева Г.Н.[a], Аширбекова Л.Ж.[a], Тажиева С.К.[a]

[a] Казахский национальный университет им. аль-Фараби, пр. аль-Фараби 71, 050040, Алматы, Казахстан;
[b] Центр устойчивого развития в Центральной Азии, Казахский национальный университет имени аль-Фараби в партнерстве с Гонконгским политехническим университетом, пр. аль-Фараби 71, 050040, Алматы, Казахстан

Для цитирования: Нурулы Е., Сансызбаева Г.Н., Аширбекова Л.Ж., Тажиева С.К. (2025). Машинное обучение в государственном управлении: систематический обзор применений, трендов и вызовов. Экономика: стратегия и практика, 20(2), 85-103, https://doi.org/10.51176/1997-9967-2025-2-85-103

АННОТАЦИЯ

Сегодня активное внедрение методов машинного обучения (далее – МО) в сферу государственного управления открывает новые возможности для прогнозирования, оценки воздействия и поддержки принятия решений, одновременно порождая целый ряд этических, институциональных и контекстуальных вызовов. Данное исследование представляет собой систематизированный обзор научных публикаций, посвящённых применению МО в государственном управлении, с акцентом на выявление ключевых тематических направлений, этических рисков и барьеров институциональной интеграции. Исходный массив данных включал 524 публикации, отобранные по целевым поисковым запросам в базах Scopus и Web of Science за период 2014–2024 гг. Фильтрация данных осуществлялась с использованием SQLite, тематическое картирование проведено в среде VOSviewer, а метаданные структурированы с помощью инструмента Elicit и последующего ручного кодирования. Анализ позволил выделить четыре функциональные области применения МО в государственном управлении: прозрачность и этика, распределение ресурсов и предоставление услуг, институциональное проектирование, а также техническая интеграция. Несмотря на достигнутый прогресс в технической реализации и повышении точности прогнозирования, во многих случаях наблюдается недостаточное внедрение механизмов обеспечения справедливости, прозрачности и участия граждан. Научная новизна работы заключается в междисциплинарном синтезе и разработке типологии институциональных вызовов, возникающих при интеграции систем МО в процессы государственного управления. Перспективы дальнейших исследований связаны с эмпирической валидацией решений, развитием методов этического аудита и институциональной готовностью к ответственному, устойчивому и контекстно адаптивному применению алгоритмических инструментов в системе государственного управления.

КЛЮЧЕВЫЕ СЛОВА: машинное обучение, государственное управление, государственная политика, внедрение технологий, стратегическое планирование, цифровая экономика

КОНФЛИКТ ИНТЕРЕСОВ: авторы заявляют об отсутствии конфликта интересов.

ФИНАНСИРОВАНИЕ: Исследование проведено в рамках грантового финансирования Комитета науки МНВО РК «Технологии искусственного интеллекта в государственном управлении: проблемы и перспективы применения» (AP23487228).

История статьи:
Получено 25 марта 2025
Принято 29 мая 2025
Опубликовано 30 июня 2025

*** Корреспондирующий автор: Нурулы Е.** – PhD докторант, старший преподаватель, старший научный сотрудник, Казахский национальный университет им. аль-Фараби, пр. аль-Фараби 71, 050040, Алматы, Казахстан, 87788797593, email: yeldar.nuruly@kaznu.edu.kz

INTRODUCTION

Machine learning (hereinafter – ML) is increasingly used in public policy to forecast demand, allocate resources, and simulate risks. Recent studies highlight its potential to enhance decision-making across healthcare, urban planning, and crisis management (Alexopoulos et al., 2019; Henman, 2020; Long & Gil-Garcia, 2023). Among the technological innovations leading this transformation is ML, a branch of artificial intelligence designed to identify patterns, predict outcomes, and generate insights without requiring explicit, rule-based programming (Ogunleye, 2024). Ethical issues such as algorithmic bias and the opacity of predictive systems remain deeply contested, particularly in fields like criminal justice and social welfare (Barn, 2020). Erroneous forecasts in such sensitive domains may inadvertently exacerbate inequality, misallocate resources, or erode public trust (Veale & Brass, 2019). Technical and infrastructural limitations, especially the inconsistent quality of administrative data and limited computational capacity in smaller jurisdictions, further complicate the effective deployment of predictive models (Khikmat et al., 2021).

A more profound concern stems from frequently omitting behavioural, cultural, and socio-economic dimensions within predictive ML applications. Without accounting for these contextual factors, models risk offering technically accurate yet socially misaligned recommendations, potentially undermining the intended policy outcomes (Cath, 2018; Huang et al., 2023; Sanchez et al., 2025). Scholars increasingly promote participatory and interdisciplinary approaches to designing and validating predictive ML systems. Co-design workshops, stakeholder consultations, and iterative feedback loops have been shown to improve the contextual fit and societal acceptance of predictive tools in public governance (Aljuneidi et al., 2023). At the same time, comparative studies across political and cultural contexts are becoming more prominent, offering insights into how local conditions mediate the success or failure of predictive interventions (Satri et al., 2024). Nevertheless, the overall scarcity of such cross-contextual research still limits the generalizability of existing findings.

Although academic interest in predictive ML for public policy has grown substantially, the research landscape exhibits thematic fragmentation and limited cross-institutional comparability. While this review initially focused on the technical and forecasting capabilities of ML, the analysis of the selected studies revealed that the most pressing challenges lie not in model architecture or prediction accuracy but in the persistent institutional, ethical, and contextual barriers hindering practical implementation. This observation necessitated a shift in analytical emphasis from technical modelling to understanding how fairness, explainability, and institutional fit determine the success or failure of ML tools in real-world governance. Accordingly, this review adopts a broader interdisciplinary lens that integrates these underexplored but critical dimensions.

Unlike prior literature reviews that predominantly focus on technical advancements or isolated policy domains, this study contributes a broader interdisciplinary synthesis by integrating ethical, institutional, and contextual dimensions of predictive ML applications. It focuses on underexplored areas such as participatory governance, cross-regional adaptability, and post-deployment accountability. This holistic perspective fills a critical gap by linking methodological practices to the realities of governance implementation.

The present study systematically reviews peer-reviewed articles from the Web of Science (WoS) and Scopus databases to address this gap. In total, 54 studies meeting predefined selection criteria were included in the analysis. The review is guided by three interrelated research questions (RQ):

RQ1: How is ML currently applied in predictive analytics for public policy?

RQ2: What emerging trends and broader implications characterize these applications?

RQ3: Which methodological practices and research gaps are most salient in the existing literature?

By synthesizing insights across these studies, the review aims to offer policymakers, scholars, and practitioners a more integrated perspective on the potentials, challenges, and ethical complexities surrounding ML-driven predictive analytics in the public sector.

LITERATURE REVIEW

Over the past two decades, the role of ML in public policy has expanded from modest administrative tools to dynamic predictive systems integrated into governance structures. Early implementations, focused primarily on structured environments, supported tasks such as fraud detection, resource allocation, risk assessment, and case prioritization using basic classification and regression algorithms (Henman, 2020; Wirtz et al., 2021). These initial applications demonstrated technical feasibility but remained constrained by narrow functional scopes and limited transparency.

The evolution of digital ecosystems, characterized by the explosion of big data, the proliferation of government digital services, and the increasing availability of real-time citizen-generated information, has significantly broadened the potential of ML in public decision-making (Alexopoulos et al., 2019; Ogunleye, 2024). This shift enabled governments to move beyond administrative efficiency toward using predictive models for more complex challenges, such as monitoring disease outbreaks, forecasting economic cycles, managing environmental risks, and anticipating infrastructural demands (Long & Gil-Garcia, 2023). Concrete applications illustrate this transformation. In urban governance, clustering algorithms have revealed localized socio-economic trends, guiding interventions in affordable housing policies and transit development (Murata, 2022). Similarly, environmental agencies have utilized ML to forecast pollution levels, climate risks, and biodiversity threats, facilitating more targeted and timely responses (Huang et al., 2023). However, deploying ML systems in the public sector has not been without significant challenges. Many early models operated as opaque "black boxes", providing limited interpretability of decision-making processes (Ridley, 2022). In domains like predictive policing, social services distribution, and tax fraud detection, the lack of transparency has triggered concerns regarding fairness, public accountability, and the amplification of historical biases embedded in datasets (Barn, 2020; Suresh & Guttag, 2021).

To address these concerns, researchers have increasingly emphasized the development of XAI frameworks, which aim to make ML model predictions understandable and auditable (Gunning et al., 2019; Papadakis et al., 2024). XAI frameworks, such as SHAP, LIME, and attention-based visualization methods, are now standard components of many governmental ML applications (Masoud, 2025). Their integration ensures that policymakers and oversight bodies can interpret algorithmic recommendations, justify decisions, and foster public trust in automated governance systems (Keller & Drake, 2021; Arora et al., 2024). ARIMA-LSTM hybrid models have demonstrated superior performance over conventional time-series models in areas such as economic forecasting and infrastructure demand projections (Dave et al., 2021).

Despite this progress, several critical challenges remain. ML models often face difficulties when transferred across different socio-political environments, primarily due to disparities in data infrastructure, human capital, legal frameworks, institutional robustness, technical capacity and effective policymaking (Khikmat et al., 2021; Sharma et al., 2022; Khan et al., 2024). Moreover, many predictive systems still inadequately integrate behavioural, cultural, personal and socio-economic factors, reducing the validity of outputs in diverse population settings (Cath, 2018; Akter et al., 2022; Sanchez et al., 2025). These gaps risk reinforcing inequalities rather than mitigating them if left unaddressed.

A pivotal trend is the rise of predictive governance, where ML models are integrated into strategic foresight frameworks to anticipate social, environmental, economic and infrastructural challenges. Governments increasingly simulate long-term policy impacts using predictive analytics to proactively address public health crises, demographic issues, economic volatility, and climate-driven migration (Rezk et al., 2018; Maffei et al., 2020; Ahern, 2025). Rather than reacting to emergent threats, policymakers aim to shape societal trajectories through data-informed decision-making preemptively.

Closely linked to this development is the expansion of multisource data integration. Advances in data aggregation now allow ML systems to integrate administrative records, sensor data, and citizen inputs, expanding their forecasting capacity for complex urban and environmental planning (Gamage, 2016; Munné, 2016; Ogunleye, 2024; Zang & You, 2023; Zhang et al., 2022). These innovations substantially enhance the predictive capacity of ML systems, allowing them to capture complex, interdependent variables that traditional models could not address. In urban planning, for example, the fusion of traffic sensor data with social media sentiment analysis has improved forecasting for infrastructure demand (Iftikhar & Khan, 2020; Long & Gil-Garcia, 2023; Fadhel et al., 2024; Qiu & Zhao, 2025).

Parallel to these technical shifts, ethical governance frameworks are increasingly institutionalized. Governments and international bodies have introduced comprehensive AI ethics guidelines mandating bias audits, fairness evaluations, and dynamic privacy assessments throughout the ML system lifecycle (Ayling & Chapman, 2021; Criado et al., 2024; Krijger, 2024; Madan & Ashok, 2023; Vatamanu & Tofan, 2025). These frameworks emphasize that predictive analytics must align with fundamental human rights, data sovereignty principles, and the evolving standards of democratic accountability (Floridi & Cowls, 2022).

Another salient trend is the growing role of citizen participation in the design, evaluation, and governance of predictive models. Citizen juries, participatory algorithm audits, and co-design initiatives are emerging as important mechanisms to ensure that ML systems reflect societal priorities and cultural nuances (Aljuneidi et al., 2023; Lahdili et al., 2024). Researchers emphasized that predictive models must be contextually adaptable to avoid failures when applied across regions with distinct governance structures, socio-economic conditions, and data infrastructures (Veale & Binns, 2017; Žliobaitė, 2017). Incorporating citizen feedback, facilitating co-design workshops, and promoting participatory evaluations are strategies being implemented to ensure that predictive systems respect societal values, accommodate diverse needs, and enhance legitimacy (Guerreiro et al., 2024; Bono Rossello et al., 2025).

The risk of reproducing entrenched social inequalities lies at the heart of ethical concerns. Predictive systems trained on historical administrative data often reflect disparities rooted in past institutional practices (Mehrabi et al., 2021). When applied to areas like law enforcement or healthcare eligibility assessments, these systems may reinforce rather than rectify structural injustices. Furthermore, the inherent opacity of many ML models, and intense neural networks obstructs critical scrutiny, making it difficult for both policymakers and affected citizens to understand or contest the logic behind algorithmic outcomes (Raji et al., 2020; Ridley, 2022). Although XAI methods are increasingly being integrated into governance frameworks, their adoption remains uneven and often superficial, especially in complex real-world deployments (Papadakis et al., 2024).

Technical limitations compound these ethical risks. Public sector datasets frequently suffer from issues such as poor coverage, lack of standardization, and temporal mismatch with evolving societal conditions (Veale et al., 2018; Otley et al., 2021). When applied outside their initial design contexts, predictive models trained under these conditions are vulnerable to brittleness. Research has shown that models optimized for high-income urban environments, for example, may perform poorly in rural or economically disadvantaged regions, where data patterns differ markedly (Žliobaitė, 2017; Guerreiro et al., 2024;). Technical fixes, such as better feature engineering or model tuning, cannot fully resolve these structural data problems deeply embedded in institutional realities.

On an operational level, public agencies often lack the technical expertise and organizational agility required to deploy and maintain ML systems responsibly (Khan et al., 2024; Wirtz & Müller, 2019). Procurement processes tend to favour proprietary solutions with limited transparency, while post-deployment auditing and recalibration practices are inconsistently applied, if at all (Leslie, 2019). Without clear accountability frameworks, the risk grows that algorithmic errors will go undetected or uncorrected, particularly in politically sensitive domains. Moreover, reliance on external vendors for critical ML infrastructure can weaken the governmental capacity to govern and independently adapt these systems over time.

Importantly, these challenges are not isolated. Ethical lapses often originate in technical flaws, such as biased or incomplete training data. Similarly, operational weaknesses–including lack of audibility, can magnify the risks posed by opaque or poorly validated models. Addressing one dimension without simultaneously considering the others is unlikely to yield sustainable improvements. Recent proposals emphasize the need for integrated governance models that embed ethical auditing, technical validation, and participatory oversight into the entire ML lifecycle (Arnstein, 2019; Jobin et al., 2019). However, scaling such frameworks from pilot initiatives to routine practice remains formidable. Achieving this goal requires better technical tools and a reconfiguration of institutional norms around accountability, expertise, and public trust.

To contextualize the evolution of ML in governance, Table 1 presents a synthesized overview of the key domains where ML applications have been observed in the reviewed literature.

Table 1. Key domains of ML application in public policy (based on reviewed studies)

Domain	Typical use cases	Example studies	ML methods employed
Economic Policy	Forecasting GDP, inflation, and public expenditures; optimizing resource allocation	(Dave et al., 2021; Osman & Muse, 2024)	ARIMA-LSTM, Random Forest (RF), Linear Models
Healthcare	Predicting hospital admissions and crisis response needs	(Guerreiro et al., 2024; Khan et al., 2024; Zang & You, 2023)	Gradient Boosting, LSTM, XGBoost
Environmental Policy	Air pollution forecasting; disaster risk modelling; climate change simulation	(Huang et al., 2023)	SVM, Decision Trees, Neural Networks

Urban Planning	Infrastructure and traffic prediction; smart city management	(Murata, 2022; Otley et al., 2021)	Clustering (k-means, DB-SCAN), LSTM
Social Services	Welfare eligibility, fraud detection, and resource targeting	(Barn, 2020; Raji et al., 2020; Suresh & Guttag, 2021)	SVM, Logistic Regression, Decision Trees
Digital Governance/ E-Gov	Enhancing transparency, automating workflows, and increasing citizen engagement	(Arora et al., 2024; Ridley, 2022)	NLP models, XAI (SHAP, LIME)

Note: compiled by the authors

Addressing these gaps requires interdisciplinary research efforts that bridge technical, ethical, and institutional dimensions. Future studies should prioritize empirical evaluations in real-world governance settings, develop measurable ethical auditing tools, investigate cross-context model adaptation, and design scalable frameworks for democratic oversight of predictive systems.

To consolidate and visualize the main insights from the reviewed studies, their methodological characteristics, practical applications, and policy implications, Appendix 1 provides a structured synthesis.

METHODOLOGY

The research used systematic methods to find peer-reviewed studies about ML applications for forecasting, impact evaluation, and decision-making support in public governance environments. WoS and Scopus served as primary data sources because they extensively cover interdisciplinary publications between policy and informatics and computational social science fields. The search query used Boolean operators to find studies where ML, policy relevance and predictive modelling converged: (“machine learning” OR “predictive analytics”) AND (“public policy” OR “government” OR “public sector”) AND (“decision-making” OR “forecasting” OR “impact”). The research query targeted three sections of data: titles, abstracts and keywords. The research focused on English-language studies published between 2014 and 2024. The selected period corresponds to the increasing popularity of AI technologies throughout government operations from 2014 to 2024. The combined query returned 524 records, 213 of which came from WoS and 311 from Scopus (see Figure 1).

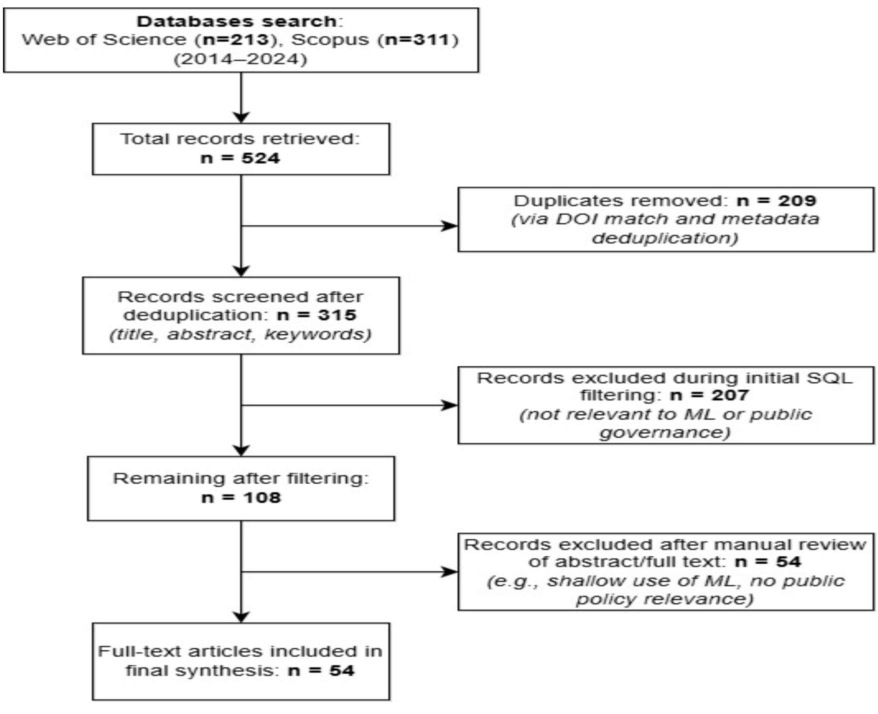

Figure 1. Flow diagram of study identification and screening

The research implemented a two-stage filtering process. The first step used structured SQL queries operated by SQLite as a lightweight relational database to remove duplicates while enforcing topical restrictions. Research studies focusing exclusively on technical subfields like precision agriculture, biomedical diagnostics and mechanical engineering were excluded if they failed to show connections to governance or policy implementation. The research included only documents with structured metadata, including abstracts, DOIs, keywords, and publication source information.

Structured SQL queries using SQLite were used for filtering, deduplication, and classification (see Appendix 2 for details).

The second stage of the selection process consisted of a manual review of abstracts, followed by a full-text evaluation to determine conceptual relevance. Studies were included if they presented original research involving the use of ML for forecasting or decision-making support, addressed domains such as budgeting, social service delivery, infrastructure, or regulation, were published in peer-reviewed outlets (including journals, conference proceedings, or academic volumes), and demonstrated analytical depth beyond superficial keyword mentions.

The study excluded articles when they consisted of opinion pieces focused exclusively on algorithmic design without decision-making relevance for the public sector. While the final corpus comprises 54 studies, this number reflects a narrow and conceptually rigorous selection process. The inclusion criteria were designed to exclude generic or technically isolated ML research and instead emphasize studies that explicitly operationalize predictive analytics in public governance contexts–linking algorithmic methods to budgeting, service delivery, regulation, or strategic decision-making. Given the interdisciplinary nature of the topic and the scoping review design, this corpus is sufficiently comprehensive to capture the dominant trends, methodological approaches, and conceptual gaps relevant to the field.

All retrieved records were exported in CSV format and cleaned using a combination of Excel and SQLite. Essential fields retained included authorship, title, abstract, keywords, DOI, link, citation count, source type, and publisher. SQLite was selected as the primary processing environment due to its transparency and ability to support precise, replicable filtering operations (Allen & Owens, 2010; Feiler, 2015). This procedural traceability ensures full reproducibility of the selection workflow. The cleaned dataset was the foundation for subsequent qualitative coding and metadata analysis.

The final corpus of 54 studies was synthesized using a combination of manual coding and digital tools. Elicit, an AI-based research assistant, supported the extraction of metadata and thematic structuring (Whitfield & Hofmann, 2023). Outputs were manually reviewed, as Elicit primarily analyses abstracts and may not capture argument depth or contextual nuance (Ejjami, 2024). VOSviewer was used to visualize co-occurrence among keywords to map thematic convergence. These networks revealed clusters like XAI in regulation, predictive service delivery, citizen-involved ML systems, and infrastructure risk forecasting. While the initial inclusion criteria emphasized predictive applications, the thematic coding revealed that institutional, ethical, and contextual integration challenges emerged more frequently than strictly technical modelling concerns. Therefore, the analytical lens of the review was broadened to reflect this empirical pattern, capturing governance-related dynamics and ethical oversight as primary themes in the synthesis. While VOSviewer is limited in capturing semantic argumentation, its utility in bibliometric mapping is well-established (Spillias et al., 2024; van Eck & Waltman, 2010).

Thus, the research methodology is based on a strict two-stage source selection procedure and a combination of quantitative and qualitative analysis methods. The use of structured SQL queries, manual peer review, as well as modern tools such as Elicit and VOSviewer ensured the representativeness and analytical depth of the selected corpus. Thanks to the integration of formal selection criteria and visual thematic mapping, the study makes it possible to identify not only technological trends but also institutional, ethical and contextual aspects of the introduction of machine learning in the field of public administration.

RESULTS

After selecting and structuring the final corpus of 54 peer-reviewed publications, a bibliometric analysis was conducted to identify the quantitative and temporal characteristics of the development of the research area. Attention was paid to the chronological distribution of publications, as it allows us to assess the degree of stability and dynamics of scientific interest in applying machine learning methods in public administration. The analysis of time trends is an important element of the methodological part, as it reflects not only the growth of academic interest, but also periods of increased practical implementation of digital solutions in the public sector.

Figure 2 shows how the number of relevant publications changed between 2014 and 2024.

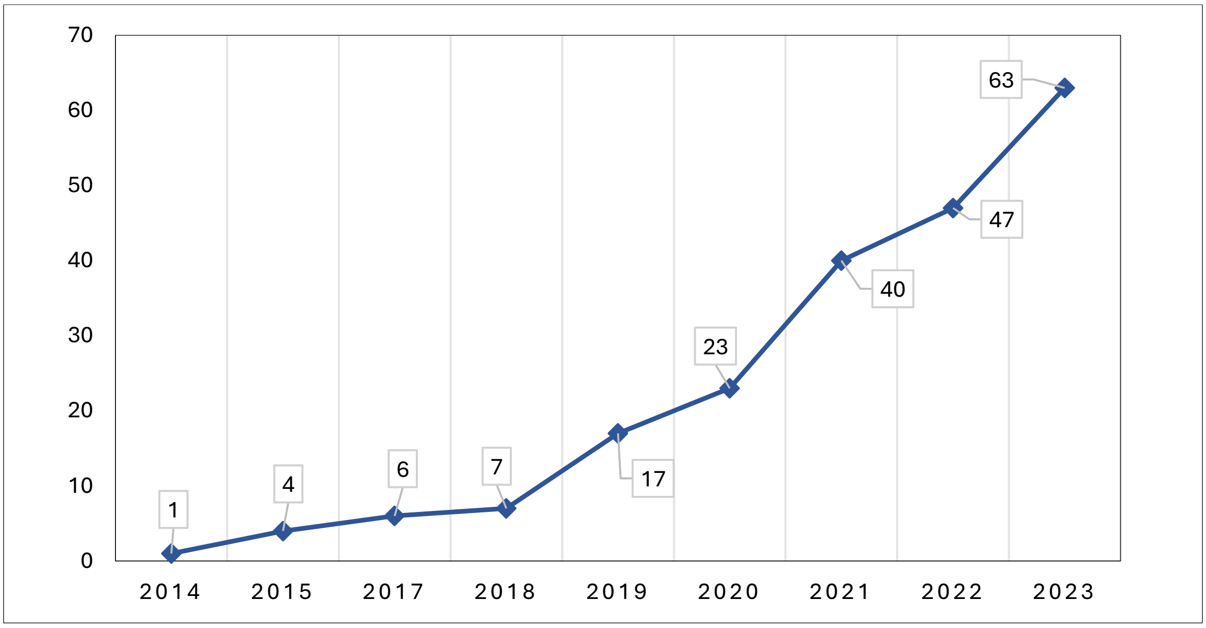

Figure 2. Number of publications by year for 2014-2024

The dynamics demonstrate a steady increase in scientific papers on applying machine learning in public administration. The most significant number of publications is in 2020, reflecting increased digital initiatives in response to the pandemic and increased interest in predictive technologies in public policy. Next, Figure 3 shows the distribution of included studies by publication type.

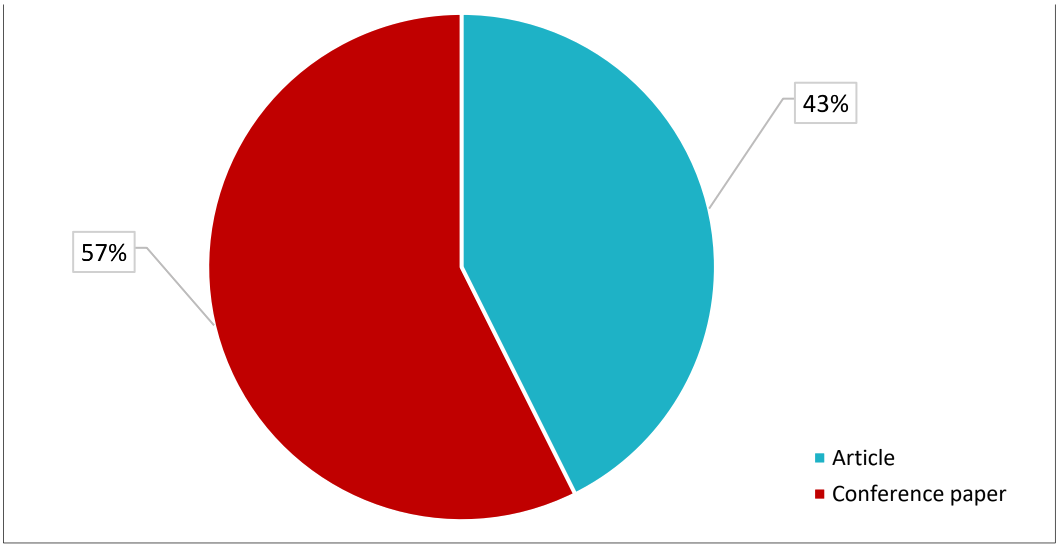

Figure 3. Distribution of included studies by document type

The scheme draws from 57% of conference papers, while journal articles comprise 43% of the total studies. The field of computer science maintains its methodological roots because conferences are the primary channels for quick research dissemination. Most present research exists in exploratory or prototyping stages because few studies achieve complete journal publication and policy-oriented reflection.

A breakdown by publisher further illustrates the dominance of technical venues (Figure 4).

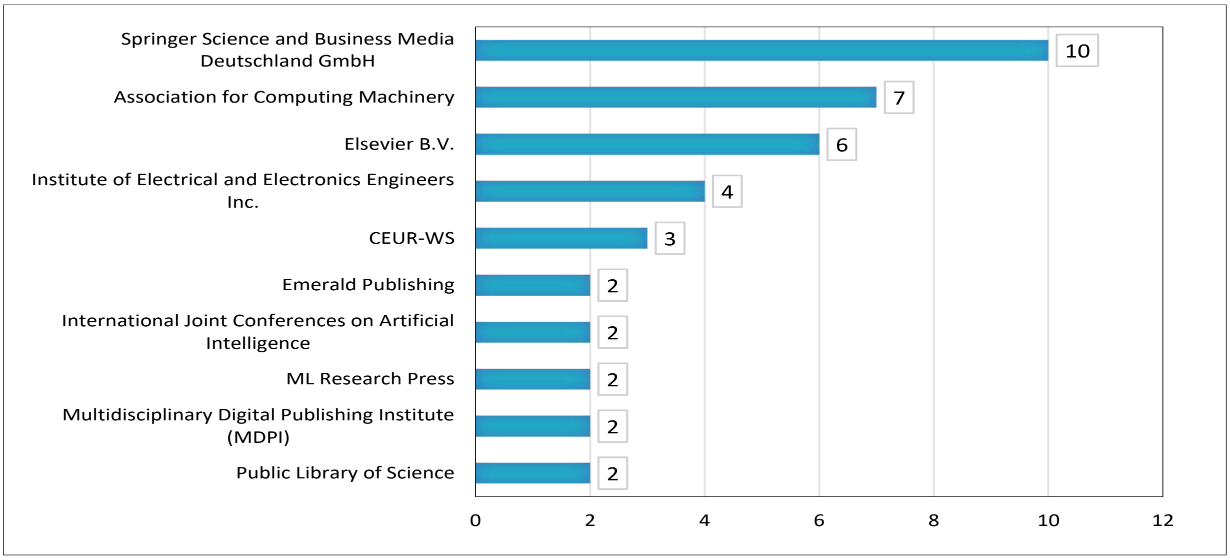

Figure 4. Distribution of studies by publishing venue

Springer was the most frequent outlet (10 studies), followed by ACM (7), Elsevier (6), and IEEE (4). Combined, these four publishers account for exactly half of the corpus. Other outlets such as CEUR-WS, MDPI, PLOS, and IJCAI occasionally appear but rarely exceed one or two studies. This highlights that research on machine learning in public policy is more often published in technical contexts and much less often in specialized social science publications. This trend may slow down the introduction of scientific developments into public administration practice, as it creates a gap between technology developers and the political and administrative community.

The RQ1 receives its answer by studying how ML applications spread across various public governance fields. The thematic coding of 54 studies produced 14 different ML application areas, summarized in Figure 5.

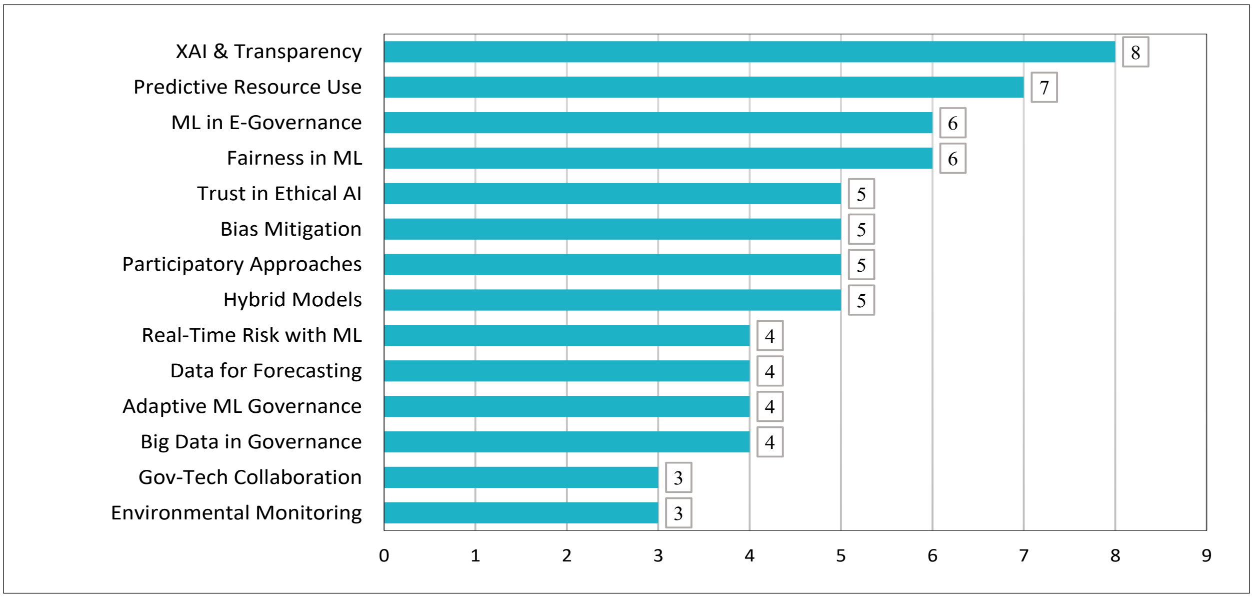

Figure 5. Frequency of ML applications in governance domains

These areas were grouped into four functional domains to facilitate comparative analysis: (1) explainability and ethics, (2) resource allocation and service delivery, (3) governance design and experimentation, and (4) technical integration. This typology is presented in Table 3.

Table 3. Functional typology of ML application areas in public governance

Functional Domain	Application areas (with frequency)	Key observations
Explainability and Ethics	Explainability & transparency (8), Fairness (6), Ethical trust (5), Bias mitigation (5)	Explainability and transparency received eight mentions, but XAI and fairness metrics have not yet received formal implementation. The discussion about ethical issues occurs primarily at theoretical levels rather than through practical implementation
Resource Allocation & Service Delivery	Predictive resource use (7), E-governance (6), Real-time risk management (4), Social forecasting (4)	The applications focus on efficiency improvement through automation and cost reduction during budgeting and, risk planning, and service delivery. Most models exist only in prototype form
Governance Design & Experimentation	Participatory approaches (5), Long-term planning (2), Infrastructure resilience (2), Adaptive governance (4)	The research areas demonstrate institutional integration alongside democratic accountability, yet they lack substantial practical implementation
Technical Integration	Big data integration (4), Hybrid modelling (5), GovTech collaboration (3), Environmental monitoring (3)	The system requires interoperability with existing infrastructure. The research demonstrates its application in experimental scenarios but lacks deployment in real-world settings

Note: compiled by the authors

The two most significant domains that emerged were "explainability and ethics". The research documents establish what should be mandatory for ML systems to operate as interpretable systems which are also socially legitimate within public sector contexts. The discourse about algorithmic fairness spreads widely yet practical methodological approaches like formal XAI techniques and fairness-aware optimization techniques are used infrequently. The three remaining domains show similar patterns since "resource allocation and service delivery" lead to the most applied work. However, the approaches stay limited to methodological specifics. At the same time, governance design and experimentation offer promising but underdeveloped methods for democratic accountability, and institutional innovation and technical integration continue to face ongoing infrastructural and interoperability challenges. The examined body demonstrates active experimentation but lacks institutional adoption and field-testing evidence.

The RQ2 synthesizing recent developments, persistent barriers, and underrepresented areas in the current literature. While interest in ML for public governance has expanded, the thematic and methodological landscape remains fragmented. One major trend is the increasing integration of predictive analytics into public service delivery. Many studies use ML to forecast resource demand, identify inefficiencies, or automate administrative decisions. However, most implementations remain experimental, with limited deployment in operational workflows or formal policy cycles. There is also a growing focus on explainability, though few studies adopt formal XAI techniques or evaluate transparency in user-facing systems. Ethical principles such as fairness, bias mitigation, and legitimacy–are often mentioned yet rarely operationalized. These patterns suggest that normative considerations are acknowledged but insufficiently embedded into modelling practice or evaluation.

Beyond technical and ethical aspects, a broader challenge concerns the limited attention to institutional learning and participatory design. Very few studies involve stakeholders in model development, test tools in real-world governance settings, or assess long-term policy impact. The literature is dominated by technically oriented contributions, with relatively few efforts to connect ML with public values, administrative constraints, or democratic accountability.

As shown in Figure 6, the key concepts related to the application of ML in public administration from several thematic clusters.

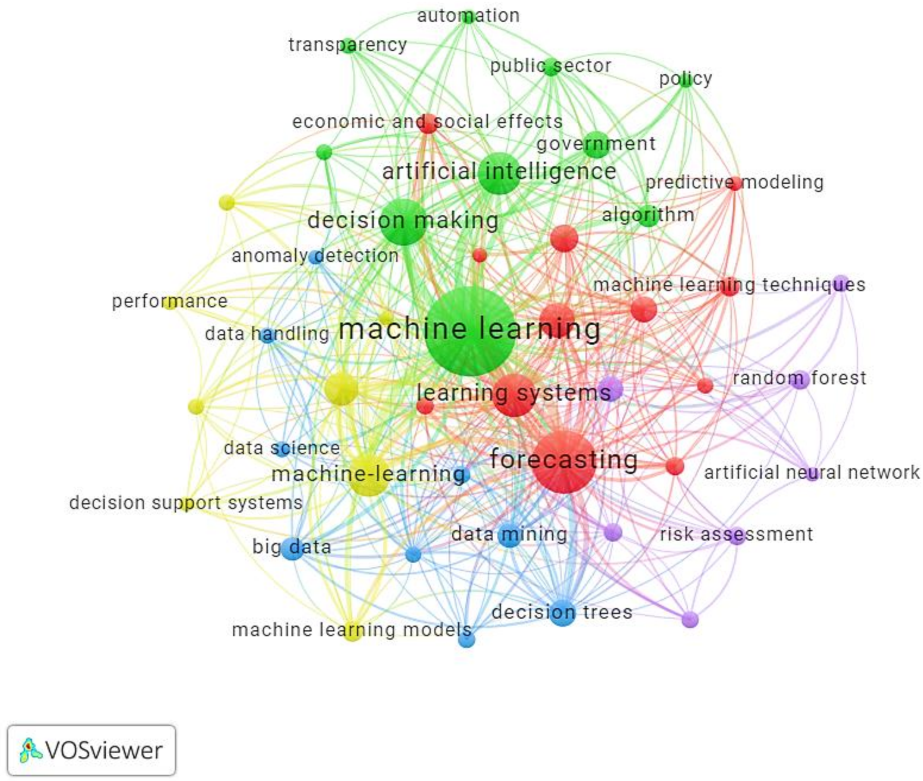

Figure 6. Keyword co-occurrence map (VOSviewer output)

Scheme 6 visualizes the thematic structure of the research field by analysing the frequency of keywords. These patterns are reflected in the bibliometric keyword analysis, which reveals a strong clustering around forecasting, decision support, and data systems–while terms like citizen engagement or co-design remain peripheral. Each cluster on the graph represents a group of concepts often found in a single publication. Table 4 summarizes the conceptual clusters identified through co-occurrence analysis.

Table 4. Conceptual clusters of ML research in public governance

Cluster	Title	Description
1 (Green)	Governance and Decision-Making	Focus on transparency, AI legitimacy, and automated systems in policymaking.
2 (Red)	Forecasting and Risk Modelling	Use of ML for crisis planning, resource allocation, and operational efficiency.
3 (Yellow)	Data Handling and Support Tools	Emphasis on data infrastructure, performance metrics, and decision support systems.
4 (Blue)	ML Models and Algorithms	Evaluation of algorithms and technical benchmarking.
5 (Purple)	Advanced ML Techniques	Specialized models like neural networks and random forests are often disconnected from policy applications.

Note: compiled by the authors

The bibliometric clustering reinforces the observed thematic fragmentation: while Cluster 1 reflects normative discourses (transparency, governance), Clusters 2 and 3 dominate in technical focus. Clusters 4 and 5, which centre on model design, appear disconnected from participatory or institutional concerns. The absence of links between Clusters 1 and 5 suggests a structural gap between algorithmic innovation and democratic accountability. Together, these findings indicate that ML research in governance is advancing rapidly in technical terms but lacks methodological diversity and institutional anchoring. Without stronger interdisciplinary collaboration and greater engagement with policy environments, many tools risk remaining academic experiments rather than scalable public-sector solutions.

The analysis of ML applications in public governance follows RQ3 to examine their methodological execution and institutional embedding based on reviewed literature. The analysis includes a review of model validation approaches, explainability methods, fairness assessment procedures, stakeholder involvement practices, and regulatory standards.

Across the corpus, methodological integration remains highly uneven. While 44 out of 54 studies provide performance metrics such as accuracy or RMSE, only a minority offer robust validation strategies. For example, stakeholder-informed model calibration or scenario-based testing appears in fewer than 15% of cases. Most evaluations are limited to internal statistical validation, with no examination of institutional fit or user impact.

Explainability receives high conceptual attention: 11 studies mention the importance of transparency, yet only four deploy formal XAI tools like SHAP or LIME. Six studies discuss fairness in rhetorical terms but none of them use fairness-aware algorithms together with mitigation pipelines. The documentation of ethical principles occurs frequently, yet the process of converting these principles into operational design or evaluation remains restricted.

The institutional integration practices demonstrate similar underdevelopment in the field. Only nine studies address regulatory issues, typically at a general level without linking to policy frameworks or legal mandates. The two studies proposed formal institutional oversight mechanisms (e.g., auditability and redress systems), but they exist only in conceptual form. Participatory design or citizen validation appears in just five studies which use post-hoc user surveys as their main method rather than co-production or real-time governance trials.

Table 5 categorizes the core integration dimensions according to frequency and typical implementation to consolidate these findings based on the 54-study corpus. Frequency levels were assigned using the following thresholds: high (>60%), moderate (20–60%), low (5–20%), and very low (<5%).

Table 5. Methodological and institutional integration dimensions in ML-for-governance studies

Integration Dimension	Frequency	Typical implementation
Performance Evaluation	High	Accuracy, precision, error rates; no contextual validation
Explainability / XAI	Moderate	Mostly SHAP/LIME or descriptive references
Fairness Metrics	Low	Conceptual only; no algorithmic debiasing or fairness auditing
Institutional Regulation	Moderate	General policy references; limited legal or procedural links
Public Trust & Legitimacy	Low	Survey-based trust proxies; no participatory mechanisms
Governance Oversight	Very Low	Framework proposals without implementation or pilot validation

Note: compiled by the authors

The research evidence demonstrates a significant gap between algorithm development and institutional implementation practices. The widespread use of technical experimentation alongside performance metric reporting exists alongside a scarcity of deep integration with governance structures and, regulatory contexts, and stakeholder processes. Most contributions maintain exploratory and conceptual approaches to ethics and institutional fit while lacking implementation trials or demonstrable impact within functioning governance systems.

DISCUSSION

The review analysed the current state of ML applications in public governance by systematically evaluating 54 peer-reviewed studies. Addressing the three interrelated RQs, the research findings demonstrate that the field is fragmented because of strong technical progress but weak institutional and ethical frameworks. Predictive analytics continues to gain popularity in public policy, but its implementation shows inconsistent patterns across different domains, methodologies and implementation settings. This study provides an original contribution by offering a structured typology of ML application domains and highlighting the key barriers–ethical ambiguity, institutional fragmentation, and contextual misalignment–that hinder their implementation in real-world governance.

ML initiatives concentrate on efficiency forecasting and risk-oriented optimization, particularly in economic and infrastructure planning. The study confirms previous research demonstrating that performance-oriented applications lead public sector digitalization during its early stages (Henman, 2020; Wirtz & Müller, 2019). The experimental nature of most prototypes indicates a difference between technological development and organizational adoption of new systems. Studies' lack of post-deployment evaluation and policy impact analysis confirms previously identified operational adoption challenges (Veale et al., 2018).

The review demonstrates that ethical awareness does not match the level of actual ethical integration in practice. Transparency and fairness are commonly used, yet operational tools such as SHAP, LIME and fairness-aware learning remain scarce. The observation supports the critical views of Raji et al. (2020) about the superficial approach to algorithmic accountability in governance settings. The lack of citizen participation stands out as a significant issue because it is absent from design principles and evaluation components despite its growing importance for legitimacy and contextual alignment (Aljuneidi et al., 2023; Lahdili et al., 2024).

The bibliometric mapping analysis shows that the field remains fragmented into separate thematic areas. The co-occurrence network shows that model benchmarking and technical optimization clusters are the most prominent. At the same time, socio-institutional themes, including co-design oversight and legal accountability, remain on the periphery. The field remains predominantly influenced by computer science paradigms, prioritising technical optimization over institutional feasibility or socio-political alignment. The field faces limitations in developing scalable context-sensitive innovations because of this misalignment, which hinders its ability to handle efficiency-equity-legitimacy trade-offs.

The review points out several promising developments despite its current limitations. First, the field now sees hybrid methodological approaches that unite ML with traditional econometric or rule-based systems as valuable connections between prediction and interpretation (Dave et al., 2021; Osman & Muse, 2024). Second, the development of modular ML frameworks continues to gain traction because they enable adaptation between jurisdictions with different legal and institutional capacities (Guerreiro et al., 2024; Žliobaitė, 2017). These trends match the current proposals for "anticipatory governance" that incorporate foresight, flexibility, and learning into the development of digital systems (Ahern, 2025).

This review identifies key patterns linking types of ML models with their implementation maturity and ethical integration, offering a structured synthesis for future research and practice. The synthesis demonstrates that sustainable ML governance must simultaneously focus on technical rigor, institutional feasibility, and democratic alignment. Moreover, the study establishes a systematic framework for future research priorities by classifying underrepresented domains, including participatory design and fairness auditing.

Policymakers should refuse to accept ML tools that only meet performance standards. They must require institutions to fit the technology, maintain transparency, and involve stakeholders at all stages of the ML lifecycle. The current procurement and auditing frameworks, along with post-deployment oversight, need redesign to establish algorithmic accountability as a built-in requirement (Arnstein, 2019; Leslie, 2019). Public managers must understand the dangers of vendor-controlled black-box solutions and develop internal capabilities for contextual adaptation model retraining and democratic supervision.

In conclusion, the current development of ML applications in public governance faces obstacles rooted in methodological silos, weak ethical frameworks, and fragmented institutional mandates. Achieving this potential requires more than improved models because it demands improved model governance systems.

CONCLUSION

This systematic review analysed 54 peer-reviewed studies to examine how ML is currently applied in public governance and to identify prevailing patterns, limitations, and directions for improvement. The findings reveal that although technical advancements in predictive analytics are considerable, the institutional integration of ML systems, their ethical alignment, and their adaptability to diverse governance contexts remain limited.

The review offers a conceptual framework that consolidates 14 areas of ML application into four broader domains: explainability and ethics, resource allocation and service delivery, governance design, and technical integration. While technical optimization dominates current implementations, integrating fairness, transparency, and accountability across domains is uneven and often underdeveloped. This discrepancy is particularly visible in the minimal use of fairness-aware approaches and the scarce involvement of citizens in the development and oversight of ML systems.

Through bibliometric mapping using VOSviewer, the analysis highlights the thematic fragmentation of the field: while algorithmic performance receives extensive attention, legal, democratic, and participatory dimensions remain peripheral. The lack of post-deployment evaluation, limited contextual adaptability, and weak institutional readiness further restrict the real-world impact of ML in governance settings. To advance responsible implementation, future research should prioritize interdisciplinary collaboration and the development of scalable ethical standards alongside institutional mechanisms that enable long-term oversight, adaptability, and public trust. Without such measures, there is a risk that ML tools will reinforce existing inequalities and undermine the legitimacy of public institutions.

Although this review provides a structured typology and analytical lens for evaluating ML in governance, it has limitations. The focus on English-language, peer-reviewed publications may have excluded relevant grey literature and non-English sources, particularly from underrepresented regions such as the Global South. Furthermore, while helpful in structuring bibliometric insights, reliance on tools like VOSviewer and Elicit may introduce bias by privileging frequently used terms and simplified abstract structures. Importantly, this review did not assess ML systems' long-term performance or policy outcomes, as such data remain scarce in the existing literature.

Future investigations should explore the longitudinal effects of ML adoption in public institutions while embedding participatory evaluation mechanisms and examining governance models across varying institutional environments. Advancing public value through ML will require sustained dialogue and cooperation between policymakers, legal scholars, computer scientists, and civil society to co-design normative standards, practical tools, and institutional safeguards.

AUTHOR CONTRIBUTIONS

Conceptualization and theory: YN; research design: YN and GS; data collection: YN, GS, LA and ST; analysis and interpretation: YN, GS, LA and ST; writing draft preparation: YN and ST; supervision: GS and LA; correction of article: YN; proofread and final approval of article: YN, GS, LA and ST. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Ahern, D. (2025). The New Anticipatory Governance Culture for Innovation: Regulatory Foresight, Regulatory Experimentation and Regulatory Learning. *European Business Organization Law Review*. https://doi.org/10.1007/s40804-025-00348-7

Akter, S., Dwivedi, Y. K., Sajib, S., Biswas, K., Bandara, R. J., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. *Journal of Business Research*, *144*, 201–216. https://doi.org/https://doi.org/10.1016/j.jbusres.2022.01.083

Alexopoulos, C., Diamantopoulou, V., Lachana, Z., Charalabidis, Y., Androutsopoulou, A., & Loutsaris, M. A. (2019). How machine learning is changing e-government. *ACM International Conference Proceeding Series*, *Part F1481*, 354–363. https://doi.org/10.1145/3326365.3326412

Aljuneidi, S., Heuten, W., Tepe, M., & Boll, S. (2023). Did that AI just Charge me a Fine? Citizens' Perceptions of AI-based Discretion in Public Administration. *ACM International Conference Proceeding Series*, 57–67. https://doi.org/10.1145/3582515.3609518

Allen, G., & Owens, M. (2010). *The Definitive Guide to SQLite* (2nd ed.). Apress. https://doi.org/10.1007/978-1-4302-3226-1

Arnstein, S. R. (2019). A Ladder of Citizen Participation. *Journal of the American Planning Association*, *85*(1), 24–34. https://doi.org/10.1080/01944363.2018.1559388

Arora, A., Gupta, M., Mehmi, S., Khanna, T., Chopra, G., Kaur, R., & Vats, P. (2024). Towards Intelligent Governance: The Role of AI in Policymaking and Decision Support for E-Governance. In *Smart Innovation, Systems and Technologies* (Vol. 379). https://doi.org/10.1007/978-981-99-8612-5_19

Ayling, J., & Chapman, A. (2021). Putting AI ethics to work: are the tools fit for purpose? *AI and Ethics*, *2*, 405–429. https://doi.org/10.1007/s43681-021-00084-x

Barn, B. S. (2020). Mapping the public debate on ethical concerns: algorithms in mainstream media. *Journal of Information, Communication and Ethics in Society*, *18*(1), 38–53. https://doi.org/10.1108/JICES-04-2019-0039

Bono Rossello, N., Simonofski, A., & Castiaux, A. (2025). Artificial intelligence for digital citizen participation: Design principles for a collective intelligence architecture. *Government Information Quarterly*, *42*(2), 102020. https://doi.org/https://doi.org/10.1016/j.giq.2025.102020

Cath, C. (2018). Governing artificial intelligence: Ethical, legal and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2133). https://doi.org/10.1098/rsta.2018.0080

Criado, J Ignacio, Sandoval-Almazán, Rodrigo, & Gil-Garcia, J Ramon. (2024). Artificial intelligence and public administration: Understanding actors, governance, and policy from micro, meso, and macro perspectives. *Public Policy and Administration*, *40*(2), 173–184. https://doi.org/10.1177/09520767241272921

Dave, E., Leonardo, A., Jeanice, M., & Hanafiah, N. (2021). Forecasting Indonesia Exports using a Hybrid Model ARIMA-LSTM. *Procedia Computer Science*, *179*, 480–487. https://doi.org/10.1016/j.procs.2021.01.031

Ejjami, R. (2024). Integrative Literature Review 5.0: Leveraging Ai and Emerging Technologies to Redefine Academic Research. *International Journal For Multidisciplinary Research*. https://doi.org/10.36948/ijfmr.2024.v06i05.28756

Fadhel, M. A., Duhaim, A. M., Saihood, A., Sewify, A., Al-Hamadani, M. N. A., Albahri, A. S., Alzubaidi, L., Gupta, A., Mirjalili, S., & Gu, Y. (2024). Comprehensive systematic review of information fusion methods in smart cities and urban environments. *Information Fusion*, *107*, 102317. https://doi.org/10.1016/J.INFFUS.2024.102317

Feiler, J. (2015). Using SQLite Basics: Storing and Retrieving Data. In J. Feiler (Ed.), *Introducing SQLite for Mobile Developers* (pp. 15–27). Apress. https://doi.org/10.1007/978-1-4842-1766-5_3

Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. In *Machine Learning and the City: Applications in Architecture and Urban Design*.

Gamage, P. (2016). New development: Leveraging 'big data' analytics in the public sector. *Public Money & Management*, *36*(5), 385–390. https://doi.org/10.1080/09540962.2016.1194087

Guerreiro, J., Garriga, R., Lozano Bagén, T., Sharma, B., Karnik, N. S., & Matić, A. (2024). Transatlantic transferability and replicability of machine-learning algorithms to predict mental health crises. *Npj Digital Medicine*, *7*(1), 227. https://doi.org/10.1038/s41746-024-01203-8

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G.-Z. (2019). XAI-Explainable artificial intelligence. *Science Robotics*, *4*(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Henman, P. (2020). Improving public services using artificial intelligence: possibilities, pitfalls, governance. *Asia Pacific Journal of Public Administration*, *42*(4), 209–221. https://doi.org/10.1080/23276665.2020.1816188

Huang, Y., Zhang, X., & Li, Y. (2023). A Novel Hybrid Model for PM2.5 Concentration Forecasting Based on Secondary Decomposition Ensemble and Weight Combination Optimization. *IEEE Access*, *11*, 119748–119765. https://doi.org/10.1109/ACCESS.2023.3327707

Iftikhar, R., & Khan, Dr. M. (2020). Social Media Big Data Analytics for Demand Forecasting: Development and Case Implementation of an Innovative Framework. *Journal of Global Information Management*, *28*, 103–120. https://doi.org/10.4018/JGIM.2020010106

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Keller, P., & Drake, A. (2021). Exclusivity and paternalism in the public governance of explainable AI. *Computer Law & Security Review*, *40*, 105490. https://doi.org/10.1016/J.CLSR.2020.105490

Khan, M. S., Umer, H., & Faruqe, F. (2024). Artificial intelligence for low income countries. *Humanities and Social Sciences Communications*, *11*(1), 1422. https://doi.org/10.1057/s41599-024-03947-w

Khikmat, R., Otabek, K., Shokhida, Y., & Khurmat, O. (2021). Developing a model and algorithm for decision support in self-government bodies using machine learning. *International Conference on Information Science and Communications Technologies: Applications, Trends and Opportunities, ICISCT 2021*. https://doi.org/10.1109/ICISCT52966.2021.9670157

Krijger, J. (2024). What About Justice and Power Imbalances? A Relational Approach to Ethical Risk Assessments for AI. *Digit. Soc.*, *3*, 56. https://doi.org/10.1007/s44206-024-00139-6

Lahdili, N., Onder, M., & Nyadera, I. (2024). Artificial Intelligence and Citizen Participation in Governance: Opportunities and Threats. *Amme Idaresi Dergisi*, *57*, 202–229.

Leslie, D. (2019). *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector* . Zenodo. https://doi.org/10.5281/zenodo.3240529

Long, Y., & Gil-Garcia, J. R. (2023). Understanding the Extent of Automation and Process Transparency Appropriate for Public Services: The AI Cases in Chinese Local Governments. *International Journal of Electronic Government Research*, *19*(1). https://doi.org/10.4018/IJEGR.322550

Madan, R., & Ashok, M. (2023). AI adoption and diffusion in public administration: A systematic literature review and future research agenda. *Government Information Quarterly*, *40*(1), 101774. https://doi.org/10.1016/J.GIQ.2022.101774

Maffei, S., Francesco, L., & and Villari, B. (2020). Data-driven anticipatory governance. Emerging scenarios in data for policy practices. *Policy Design and Practice*, *3*(2), 123-134. https://doi.org/10.1080/25741292.2020.1763896

Masoud, N. (2025). Artificial intelligence and unemployment dynamics: an econometric analysis in high-income economies. *Technological Sustainability*, *4*(1), 30–50. https://doi.org/10.1108/TECHS-04-2024-0033

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A Survey on Bias and Fairness in Machine Learning. *ACM Comput. Surv.*, *54*(6). https://doi.org/10.1145/3457607

Munné, R. (2016). *Big Data in the Public Sector*. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-21569-3_11

Murata, T. (2022). Policy Making Based on Real-Scale Social Simulations. *2022 Joint 12th International Conference on Soft Computing and Intelligent Systems and 23rd International Symposium on Advanced Intelligent Systems, SCIS and ISIS 2022*. https://doi.org/10.1109/SCISISIS55246.2022.10001860

Ogunleye, O. S. (2024). Using artificial intelligence to enhance e-government services delivery through data science and machine learning. In *Machine Learning and Data Science Techniques for Effective Government Service Delivery*. https://doi.org/10.4018/978-1-6684-9716-6.ch001

Osman, B. M., & Muse, A. M. S. (2024). Predictive analysis of Somalia's economic indicators using advanced machine learning models. *Cogent Economics and Finance*, *12*(1). https://doi.org/10.1080/23322039.2024.2426535

Otley, A., Morris, M., Newing, A., & Birkin, M. (2021). Local and application-specific geodemographics for data-led urban decision making. *Sustainability (Switzerland)*, *13*(9), 4873. https://doi.org/10.3390/su13094873

Papadakis, T., Christou, I. T., Ipektsidis, C., Soldatos, J., & Amicone, A. (2024). Explainable and transparent artificial intelligence for public policymaking. *Data & Policy*, *6*, e10. https://doi.org/10.1017/dap.2024.3

Qiu, J., & Zhao, Y. (2025). Traffic Prediction with Data Fusion and Machine Learning. *Analytics*, *4*(2). https://doi.org/10.3390/analytics4020012

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 33–44. https://doi.org/10.1145/3351095.3372873

Rezk, M. A., Ojo, A., El Khayat, G. A., & Hussein, S. (2018). A Predictive Government Decision Based on Citizen Opinions: Tools & Results. *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance*, 712–714. https://doi.org/10.1145/3209415.3209504

Ridley, M. (2022). Explainable Artificial Intelligence (XAI): Adoption and Advocacy. *Information Technology and Libraries*, *41*(2). https://doi.org/10.6017/ital.v41i2.14683

Sanchez, T. W., Brenman, M., & Ye, X. (2025). The Ethical Concerns of Artificial Intelligence in Urban Planning. *Journal of the American Planning Association*, *91*(2), 294–307. https://doi.org/10.1080/01944363.2024.2355305

Satri, J., El Mokhi, C., & Hachimi, H. (2024). Predicting the outcome of regional development projects using machine learning. *IAES International Journal of Artificial Intelligence*, *13*(1), 863–875. https://doi.org/10.11591/ijai.v13.i1.pp863-875

Sharma, M., Luthra, S., Joshi, S., & Kumar, A. (2022). Implementing challenges of artificial intelligence: Evidence from public manufacturing sector of an emerging economy. *Government Information Quarterly*, *39*(4), 101624. https://doi.org/https://doi.org/10.1016/j.giq.2021.101624

Spillias, S., Tuohy, P., Andreotta, M., Annand-Jones, R., Boschetti, F., Cvitanovic, C., Duggan, J., Fulton, E., Karcher, D., Paris, C., Shellock, R., & Trebilco, R. (2024). Human-AI collaboration to identify literature for evidence synthesis. *Cell Reports Sustainability*. https://doi.org/10.1016/j.crsus.2024.100132

Suresh, H., & Guttag, J. (2021). A framework for understanding sources of harm throughout the machine learning life cycle. *Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 1–9. https://doi.org/10.1145/3465416.3483305

van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

Vatamanu, A. F., & Tofan, M. (2025). Integrating Artificial Intelligence into Public Administration: Challenges and Vulnerabilities. *Administrative Sciences*, *15*(4). https://doi.org/10.3390/admsci15040149

Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, *4*(2), 2053951717743530. https://doi.org/10.1177/2053951717743530

Veale, M., & Brass, I. (2019). *Administration by Algorithm? Public Management Meets Public Sector Machine Learning*. Oxford Academic. https://doi.org/10.1093/oso/9780198838494.003.0006

Veale, M., Van Kleek, M., & Binns, R. (2018). Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3173574.3174014

Whitfield, S., & Hofmann, M. A. (2023). Elicit: AI literature review research assistant. *Public Services Quarterly*, *19*(3), 201-207. https://doi.org/10.1080/15228959.2023.2224125

Wirtz, B. W., & Müller, W. M. (2019). An integrated artificial intelligence framework for public management. *Public Management Review*, *21*(7), 1076–1100. https://doi.org/10.1080/14719037.2018.1549268

Wirtz, B. W., Langer, P. F., & Fenner, C. (2021). Artificial Intelligence in the Public Sector - a Research Agenda. *International Journal of Public Administration*, *44*(13), 1103-1128. https://doi.org/10.1080/01900692.2021.1947319

Zang, J., & You, P. (2023). An industrial IoT-enabled smart healthcare system using big data mining and machine learning. *Wireless Networks*, *29*(2), 909-918. https://doi.org/10.1007/s11276-022-03129-z

Zhang, D., Pee, L. G., Pan, S. L., & Cui, L. (2022). Big data analytics, resource orchestration, and digital sustainability: A case study of smart city development. *Government Information Quarterly*, *39*(1), 101626. https://doi.org/https://doi.org/10.1016/j.giq.2021.101626

Žliobaitė, I. (2017). Measuring discrimination in algorithmic decision making. *Data Mining and Knowledge Discovery*, *31*(4), 1060–1089. https://doi.org/10.1007/s10618-017-0506-1

Information about the authors

***Yeldar Nuruly** – PhD candidate, Senior Lecturer, Senior Research Fellow, Al-Farabi Kazakh National University, Almaty, Kazakhstan, email: yeldar.nuruly@kaznu.edu.kz, ORCID ID: https://orcid.org/0000-0002-9321-2285
Galiya N. Sansyzbayeva – Doc. Sc. (Econ.), Professor, Al-Farabi Kazakh National University, Almaty, Kazakhstan, email: halima.sansyzbaeva@kaznu.edu.kz, ORCID ID: https://orcid.org/0000-0002-9992-4005
Laura Z. Ashirbekova – Cand. Sc. (Econ.), Associate Professor, Al-Farabi Kazakh National University, Almaty, Kazakhstan, email: laura.ashyrbekova@kaznu.edu.kz, ORCID ID: https://orcid.org/0000-0003-0377-7854
Samal K. Tazhiyeva – Cand. Sc. (Econ.), Senior Lecturer, Al-Farabi Kazakh National University, Almaty, Kazakhstan, email: samal.tazhyeva@kaznu.edu.kz, ORCID ID: https://orcid.org/0009-0006-8148-0625

Авторлар туралы мәліметтер

***Нұрұлы Е.** – PhD докторант, аға оқытушы, аға ғылыми қызметкер, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан, email: yeldar.nuruly@kaznu.edu.kz, ORCID ID: https://orcid.org/0000-0002-9321-2285
Сансызбаева Г.Н. – э.ғ.д., профессор, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан, email: halima.sansyzbaeva@kaznu.edu.kz, ORCID ID: https://orcid.org/0000-0002-9992-4005
Аширбекова Л.Ж. – э.ғ.к., қауымдастырылған профессор, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан, email: laura.ashyrbekova@kaznu.edu.kz, ORCID ID: https://orcid.org/0000-0003-0377-7854
Тажиева С.К. – э.ғ.к., аға оқытушы, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан, email: samal.tazhyeva@kaznu.edu.kz, ORCID ID: https://orcid.org/0009-0006-8148-0625

Сведения об авторах

***Нурулы Е.** – PhD докторант, старший преподаватель, старший научный сотрудник, Казахский национальный университет им. аль-Фараби, Алматы, Казахстан, email: yeldar.nuruly@kaznu.edu.kz, ORCID ID: https://orcid.org/0000-0002-9321-2285
Сансызбаева Г.Н. – д.э.н., профессор, Казахский национальный университет им. аль-Фараби, Алматы, Казахстан, email: halima.sansyzbaeva@kaznu.edu.kz, ORCID ID: https://orcid.org/0000-0002-9992-4005
Аширбекова Л.Ж. – к.э.н., ассоциированный профессор, Казахский национальный университет им. аль-Фараби, Алматы, Казахстан, email: laura.ashyrbekova@kaznu.edu.kz, ORCID ID: https://orcid.org/0000-0003-0377-7854
Тажиева С.К. – к.э.н., старший преподаватель, Казахский национальный университет им. аль-Фараби, Алматы, Казахстан, email: samal.tazhyeva@kaznu.edu.kz, ORCID ID: https://orcid.org/0009-0006-8148-0625

Appendix 1

A comparative synthesis of ML applications in public governance (scope, methods, and integration levels)

Application area	Typical ML models	Real-world implementation level	Ethical and institutional integration	Observed gaps or tensions	Geographic focus	Representative studies
Economic Forecasting	ARIMA-LSTM, RF, Regression	High	Low	Lacks fairness metrics; no participatory oversight	Global North (Europe, OECD); Africa	(Dave et al., 2021; Osman & Muse, 2024)
Healthcare	Gradient Boosting, LSTM, SVM	Medium	Medium	Weak explainability; models rarely adapted for underserved groups	South Asia, Europe	(Arora et al., 2024; Guerreiro et al., 2024; Khan et al., 2024)
Crisis Management / Emergency Response	Time-series ML, LSTM, Simulation models	Medium	Low–Medium	Limited stakeholder input; poor contextual transferability	Global South; MENA	(Ahern, 2025; Khan et al., 2024; Rezk et al., 2018)
Environmental Monitoring	Decision Trees, Neural Networks, SVM	Medium	Low	Underrepresented in policy scenarios; ethical risk underexplored	China, EU, Brazil	(Huang et al., 2023)
Urban Planning	Clustering, k-means, LSTM	Medium	Low	Absence of co-design or validation by local authorities	Japan, UK, US	(Murata, 2022; Otley et al., 2021)
Social Services	Decision Trees, Logistic Regression	High	Low	Risk of bias amplification; black-box deployment	US, UK	(Barn, 2020; Raji et al., 2020; Suresh & Guttag, 2021)
Digital Bureaucracy	NLP, SHAP, LIME (XAI)	Low–Medium	High (in theory)	Transparency emphasized but rarely implemented in workflows	Western Europe	(Papadakis et al., 2024; Ridley, 2022)
Citizen Participation	Human-in-the-loop, Co-Design Methods	Low	High	Lacks scalability; mostly conceptual or pilot-level	Experimental (Europe)	(Aljuneidi et al., 2023; Lahdili et al., 2024)
Fairness/ Accountability	Fairness-aware ML, Causal Inference	Low	High (academic only)	Not embedded in institutional ML cycles	Primarily academic, Global North	(Floridi & Cowls, 2022; Raji et al., 2020)

**Implementation, ethical integration, and regional focus were assessed based on content reported in each study. The classification methodology is explained in Section 3.4.*

Appendix 2

SQL Query which used for filtering and categorizing articles

```sql
-- Step 1: Create a cleaned dataset with relevant fields
CREATE TABLE cleaned_articles AS
SELECT DISTINCT authors, title, abstract, keywords, DOI, link, citation_count, source, publisher, publication_year
FROM articles
WHERE
    (abstract LIKE '%machine learning%' OR keywords LIKE '%machine learning%' OR title LIKE '%machine learning%')
    AND (abstract LIKE '%predictive analytics%' OR keywords LIKE '%predictive analytics%' OR title LIKE '%predictive analytics%')
    AND (abstract LIKE '%public policy%' OR keywords LIKE '%public policy%' OR title LIKE '%public policy%'
         OR abstract LIKE '%government%' OR keywords LIKE '%government%' OR title LIKE '%government%'
         OR abstract LIKE '%public sector%' OR keywords LIKE '%public sector%' OR title LIKE '%public sector%')
    AND (abstract LIKE '%decision-making%' OR keywords LIKE '%decision-making%' OR title LIKE '%decision-making%'
         OR abstract LIKE '%forecasting%' OR keywords LIKE '%forecasting%' OR title LIKE '%forecasting%'
         OR abstract LIKE '%impact%' OR keywords LIKE '%impact%' OR title LIKE '%impact%')
    AND publication_year BETWEEN 2014 AND 2024
    AND source NOT LIKE '%agriculture%' AND source NOT LIKE '%healthcare%' AND source NOT LIKE '%social science%';

-- Step 2: Categorize articles by type
SELECT
    CASE
        WHEN source LIKE '%conference%' THEN 'Conference Paper'
        WHEN source LIKE '%journal%' THEN 'Journal Article'
        WHEN source LIKE '%book%' THEN 'Book Chapter'
        ELSE 'Other'
    END AS article_type,
    COUNT(*) AS article_count
FROM cleaned_articles
GROUP BY article_type;

-- Step 3: Count total articles and their distribution
SELECT
    COUNT(*) AS total_articles,
    COUNT(DISTINCT DOI) AS unique_articles
FROM cleaned_articles;

-- Step 4: Export a refined dataset for manual screening
SELECT
    authors, title, abstract, keywords, DOI, link, citation_count, source, publisher, publication_year
FROM cleaned_articles;
```