# Artificial Intelligence in Public Administration and Public Service Delivery: A Bibliometric Analysis and Contextual Synthesis (2020–2025)

[1] Fadilah Abdul Azis*, [2] Muhammad Nur Othman, [1] Nik Syahrim Nik Anwar, [3] Roshaliza Hamidon

[1] Faculty of Electrical Technology and Engineering, Universiti Teknikal Malaysia Melaka, Malaysia

[2] Faculty of Mechanical Technology and Engineering, Universiti Teknikal Malaysia Melaka, Malaysia

[3] Faculty of Mechanical Engineering and Technology, Universiti Malaysia Perlis, Malaysia

DOI: https://doi.org/10.47772/IJRISS.2026.101000031

Received: 03 October 2026; Accepted: 08 October 2026; Published: 28 October 2026

## ABSTRACT

Artificial intelligence (AI) is increasingly discussed in public administration, but growth in publications does not equate to successful deployment or demonstrated public value. This study maps scholarship on AI in public administration and public service delivery within a 2020 - 2025 scope. Bibliometric analysis of 112 records from Scopus and Web of Science was combined with a bounded synthesis of contextual studies. The included publications span 2021 - 2025 and received 3,035 citations; 103 were cited, and the corpus h-index was 25. Annual output increased from six publications in 2021 to 52 in 2025, with 70.54% appearing in 2024 - 2025. Indonesia and Spain each contributed ten publications, whereas the Netherlands recorded the highest citation total among the listed countries. The author-keyword network contained 26 terms across four clusters: AI-enabled public-sector activity, public administration and transparency, generative AI and policy, and digital-government transformation. Highly cited documents emphasized accountability, adoption, and service delivery. These findings locate the retrieved literature primarily within the broader public sector, particularly regarding AI service performance. Comparison with relevant reviews identifies a need to connect AI with specific administrative tasks and independently assessed community outcomes. The evidence supports a research map and an evaluation agenda, while citation counts, keyword associations, and contextual studies do not establish causal effects of AI on public-service performance or community outcomes.

**Keywords:** artificial intelligence; public administration; public services delivery; community outcomes; bibliometric analysis;

## INTRODUCTION

Artificial intelligence (AI) is increasingly incorporated into debates about how public organizations allocate resources, support administrative decisions, communicate with citizens, and deliver services. In public administration, its significance depends on the task to which an output is attached and the consequences of acting on it. Public-administration studies provide a basis for assessing these relationships. Busuioc, (2021) examines the accountability challenges of algorithmic decision-making, while Gesk & Leyer, (2022) show that citizens' acceptance varies across public-service scenarios. These perspectives connect AI adoption to institutional responsibility and service experience. Technical performance alone does not determine whether a service becomes more accessible, a procedure becomes fairer, or administrative work becomes more efficient.

The administrative setting introduces requirements beyond prediction where algorithmic decision-making creates difficulties for public accountability when institutional actors must explain and justify decisions (Busuioc, 2021). Parycek et al., (2023) situate automation within administrative procedures and identify framework conditions that shape its appropriate use. These accounts direct attention to who interprets an automated output, which responsibilities remain with public officials, and how affected people can challenge a decision. They provide a conceptual basis for examining AI systems as components of administrative arrangements rather than assessing their performance within public-service procedures.

Empirical research on broader public-sector AI also shows why adoption and outcomes require separate treatment. Organizational challenges in three Saudi public-sector organizations, including misalignment between AI and managerial decision-making, infrastructure, and data sharing (Alshahrani et al., 2022).

Meanwhile, Gesk & Leyer, (2022) find that citizens' acceptance varies across public-service scenarios. Berman et al., (2024) found difficulties with transparency and interpretability in a Swedish employment-service decision-support system. These studies concern different technologies and settings. Together, they support contextual questions about organizational readiness, citizen acceptance, and accountable service delivery, while their findings require assessment within each administrative setting.

This review examines research at the intersection of AI, public administration, and community-facing service delivery within a nominal 2020–2025 scope. Its population of interest comprises public organizations, administrative personnel, and people affected by public services. The concept concerns AI-supported administrative activity, including adoption, automation, generative AI, and governance, with attention to the distinction between task-level and community-level outcomes. The context is public administration and community service delivery. The bibliometric corpus therefore maps an adjacent administrative-AI evidence base within which these implementation and outcome questions must be interpreted.

The aim is to describe the scale, temporal development, geographic participation, citation visibility, and conceptual organization of this literature, then use relevant studies to identify defensible evaluation priorities. Bibliometric evidence establishes where research attention is concentrated. Contextual synthesis clarifies why work processes, institutional responsibility, and community consequences require different measures. The contribution is a bounded map of these relationships and their evidential limits. It does not estimate a pooled technological effect or imply that every included publication evaluates a deployed vision system.

# LITERATURE REVIEW

## 1. AI capability and the administrative task

AI systems deployed in public services can support classification, prediction, text generation, triage, or recognition, but the relevant performance criteria depend on the administrative task. Technical performance is a necessary but incomplete basis for assessing an AI system used in public work. Berman et al., (2024) relate administrative automation to procedural requirements, technological limitations, and framework conditions. Maragno et al., (2023) examine contextual factors that shape public-sector AI implementation. A public organization must additionally determine what an output means for its procedure, whether an official verifies it, and how an error affects a person. Those questions concern the administrative use of a technical output and cannot be answered by a model-performance score alone.

Parycek et al., (2023) explain that administrative automation is bounded by procedural requirements and the characteristics of the underlying technology. This perspective helps distinguish a system that assists an official from one that effectively determines an outcome. For the present review, the relevant analytical sequence is administrative input, AI output, administrative action, and community consequence. This sequence is an organizing framework for future evaluation, rather than a causal model tested by the bibliometric analysis. Evidence at one stage should not be substituted for evidence at another.

## 2. Organizational implementation and routine work

Public-sector AI research places considerable emphasis on implementation conditions. Technology–Organization–Environment framework was combined with Technology Affordances and Constraints Theory to examine how contextual factors shape AI implementation (Maragno et al., 2023). Alshahrani et al., (2022) use the attention-based view to show how situated operational attention and the structural allocation of attention interact during assimilation. The studies share an interest in the organization surrounding the technology, although their theoretical lenses emphasize different relationships.

These perspectives suggest that task integration is an empirical question. A model may be available even when

staff lack appropriate procedures, usable explanations, or the authority to respond to uncertain outputs. Berman et al. (2024) examine this difficulty in public employment services, where transparency and stakeholder involvement complicate the ambition to deliver trustworthy AI. For AI-enabled administrative workflows, these findings motivate investigation of review practices, staff training, escalation rules, and the handling of false matches. They do not demonstrate the prevalence of those practices in the present corpus or prove that any single organizational arrangement produces better outcomes.

## 3. Service efficiency, acceptance, and community outcomes

Studies of public-sector AI use several outcome concepts that should remain distinct. Kulal et al., (2024) examine public-service delivery through responses from chief information officers in India. Their findings concern reported improvements and organizational readiness. Kleiman & Barbosa, (2025) assess initial experiences with a language-model chatbot in Brazil's federal public sector and describe the prospect of reducing program-related response time. Alamsyah & Aryfiyanto, (2025) combine a systematic review with expert interviews to propose a generative-AI governance framework for Indonesian public services. These studies address perceived benefits, early service use, and proposed implementation arrangements; they do not constitute a controlled evaluation of AI's impact on the community.

Citizen acceptance is another outcome. (Gesk & Leyer, 2022) analyze 329 participants across six public-service scenarios and find differences between general and specific services. Acceptance therefore cannot be assumed from technological availability or an organizational efficiency claim. Rannisto & Vainionpää, (2025) examine citizens' everyday trust in algorithmized public services in Finland, extending evaluation to the experiences through which people relate to these services. Moon, (2023) develops a participatory-governance argument for inclusive AI and social good. Together, these contributions suggest an evaluation structure in which staff experience, citizen acceptance, service performance, and distributive consequences are recorded separately. Convergence across those measures would need to be demonstrated, particularly where an application serves people with different needs or resources.

## 4. Accountability and the evidential limits of deployment

Responsibility persists when public bodies use algorithms. In a preregistered survey experiment with 2,483 respondents in the Netherlands, Alon-Barkat et al., (2025) find that public organizations are not held to a lower responsibility standard for algorithmic than for human discrimination. Respondents assigned more responsibility when the algorithm was developed internally. The result concerns the attribution of responsibility in the tested scenarios, rather than the frequency of discrimination in operational systems. It nevertheless complements Busuioc, (2021) conceptual analysis by showing why the institutional location of automated decisions matters to citizens.

Governance research also extends beyond individual applications. Sienkiewicz-Małyjurek & Zyzak, (2025) synthesize 59 publications on digital technologies in public-administration networks and identify fragmented research on their role in collaboration. Ooko, (2025) proposes a decolonial communication framework for AI in African public services, with attention to language and cultural context. The former supplies a systematic account of network-level research; the latter is a conceptual research note rather than a test of improved outcomes. These contributions broaden the basis for governance research into how administrative arrangements and communication shape the conditions under which AI services are used.

## 5. Research Gaps and Future Directions

Table 1 presents three reviews directly related to the present administrative and community focus. Caiza et al., (2024) examine AI in government decision-making; Nemavhola et al., (2025) concentrate on technical face-recognition research; and Stiernströmer, (2026) links vision technology to a public-service setting but identifies limited operational evidence, with greater emphasis on public perceptions. Their different boundaries suggest a need to connect AI capabilities, the administrative tasks they support, and the community outcomes by which their use is judged. The two facial-recognition reviews provide domain-specific comparators, while government decision-making supplies the broader administrative context.

The resulting gap concerns integration and evaluation across public-sector AI research. Comparing these selected reviews does not establish that no equivalent review exists. The present analysis characterizes the broader retrieved literature through publication, citation, country, three-field, and keyword evidence.

Contextual studies then distinguish organizational implementation, service acceptance, and public accountability from technical performance. This approach connects adoption, administrative automation, generative AI, and governance within a common research map, while retaining the differences between application settings and outcome measures. The research questions address the observable bibliometric structure and use it to delimit, rather than overstate, the contribution to understanding administrative tasks and community outcomes.

**Table 1. Comparison of three related bibliometric studies.**

| Authors (Year) | Objective (Summary) | Data Source & Year Coverage | TDE | Bibliometric Attributes Examined |
|---|---|---|---|---|
| Caiza et al., (2024) | Examine recent trends, advancements, benefits, challenges, and implications of AI integration in governmental decision-making and public-service provision. | Scopus, SpringerLink, Web of Science and complementary academic sources; **2019–2023**. | **43** | AI applications in government decision-making; administrative efficiency and effectiveness; public-policy development; regulatory frameworks; transparency; accountability; public acceptance; ethical and implementation challenges; keyword co-occurrence and international research collaboration. |
| Nemavhola et al., (2025) | Map the literature on deep-learning and CNN techniques for face recognition, with particular attention to image/video data, occlusion, methodological approaches, and research gaps. | **12 academic databases**, including ACM Digital Library, IEEE Xplore, Springer, Scopus, ScienceDirect, ProQuest and Taylor & Francis; final analysis covers studies published **2013–2023**. | **266** | Publication trends and countries; research aims; image versus video data; clean versus occluded data; supervised/unsupervised learning; real-time usability; CNN architectures; recognition challenges including occlusion, distance, camera angle and lighting; methodological and technical research gaps. |
| Stiernströmer, (2026) | Map empirical evidence on facial recognition technology in law enforcement, emphasizing public perceptions, operational outcomes, legitimacy, accountability, effectiveness, and research gaps. | Scopus, Web of Science, PsycInfo, Criminology Collection and Google Scholar; **no search-year restriction**; included studies were published **2020–2024**. | **11** | Study design and geographic context; theoretical frameworks; policing context; population focus; regulatory environment; FRT characteristics; public acceptance, privacy, trust and legitimacy; deployment and oversight; racial disparities; crime and investigative outcomes; operational effectiveness and research gaps. |

**Note. TDE = Total Documents Examined, referring to each study's final analyzed dataset. Publication year is distinct from the period covered by the dataset.**

## RESEARCH QUESTIONS

The review is organized around six questions that correspond to the available analyses:

RQ1: What publication, authorship, and citation profile characterizes the retrieved literature?

RQ2: How do publication output and accumulated citations vary across publication years?

RQ3: Which countries contribute to the corpus, and how do publication and citation patterns differ among them?

RQ4: Which documents attract the most citations, and what does their subject matter indicate about the scope of the retrieved literature?

RQ5: Which country-author-keyword relationships are represented in the three - field analysis?

RQ6: Which conceptual groups and connections are represented by author - keyword co - occurrence?

These descriptive questions support the broader scoping objective of locating relevant research and identifying what additional evidence is needed to evaluate AI-enabled administrative workflows and community outcomes.

# METHODS

## 1. Review scope and search strategy

This study employed a bibliometric approach complemented by contextual synthesis to examine research trends, publication characteristics, and thematic developments in artificial intelligence (AI) within public administration and public service delivery. Bibliographic records were retrieved from Scopus and the Web of Science (WoS) Core Collection, selected for their extensive coverage of peer-reviewed scientific literature and availability of bibliometric metadata. Both database searches were conducted on September 15, 2026, covering publications from 2021 to 2025.

The search strategy was developed around three interconnected conceptual dimensions: (i) artificial intelligence and related technologies, (ii) administrative and governance activities, and (iii) public services and community-related outcomes. Boolean operators (AND and OR) were used to combine the search terms, enabling the identification of publications examining AI applications in administrative, organizational, and public service contexts. The search was intentionally broad, encompassing general AI technologies rather than restricting the analysis to specific applications such as computer vision or machine vision.

The searches were restricted to the Social Sciences subject area, English-language publications, selected scholarly document types, and open-access publications. Scopus searches were conducted using the TITLE-ABS-KEY field, whereas WoS records were retrieved through its Smart Search interface. The complete database-specific search expressions, search dates, and applied restrictions are presented in Table 2 to support methodological transparency and reproducibility. The implications of restricting the dataset to open-access publications are acknowledged as a potential source of selection bias.

Following database retrieval, 104 records were obtained from Scopus and 13 from WoS. Five overlapping records were identified during database merging, resulting in a consolidated dataset of 112 unique publications. The merged records comprised 99 Scopus-only publications, eight WoS-only publications, and five publications indexed in both databases. The bibliometric dataset was subsequently processed using BiblioSpy® for descriptive performance analysis and bibliometric network visualization. The study identification and selection process is illustrated in Figure 1.

Table 2. Search strings used for the bibliometric review process

| Database | Search date | Boolean Operator Used | Record retrieved |
|---|---|---|---|
| **SCOPUS** | 15-Sep-26 | TITLE-ABS-KEY ( ( "artificial intelligence" OR "ai" OR "machine vision" OR "computer vision" ) AND ( "administration" OR | 104 |

| | | | |
|---|---|---|---|
| | | "management" OR "governance" OR "leadership" ) AND ( "community outcomes" OR "social impact" OR "public services" OR "community development" ) ) AND ( LIMIT-TO ( SUBJAREA , "SOCI" ) ) AND ( LIMIT-TO ( DOCTYPE , "ar" ) OR LIMIT-TO ( DOCTYPE , "cp" ) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) ) AND ( LIMIT-TO ( PUBSTAGE , "final" ) ) AND ( LIMIT-TO ( OA , "all" ) ) AND PUBYEAR > 2020 AND PUBYEAR < 2026 | |
| **Web of Science Core Collection** | 15-Sep-26 | ("artificial intelligence" OR "ai" OR "machine vision" OR "computer vision") AND ("administration" OR "management" OR "governance" OR "leadership") AND ("community outcomes" OR "social impact" OR "public services" OR "community development") and 2025 or 2024 or 2023 or 2022 or 2021 (Publication Years) and Social Sciences Other Topics (Research Areas) and Article or Proceeding Paper or Review Article (Document Types) and English (Languages) and All Open Access (Open Access) | 13 |

## 2. Source selection and evidence boundaries

Fig. 1 shows the flow of screening order using PRISMA. It identifies 904 Scopus and 253 Web of Science records, totaling 1,157, and records 401 at the screening stage. It lists restrictions related to Social Sciences, article or conference paper status, English, final publication stage, and open access. The later selection sequence comprises 180 reports sought for retrieval, 63 not open access, 117 assessed for eligibility, and five duplicate removed, leaving 112. The analysis was restricted to open-access publications to facilitate consistent access to full-text articles for the contextual synthesis component. These final stages reconcile with the analytical dataset.

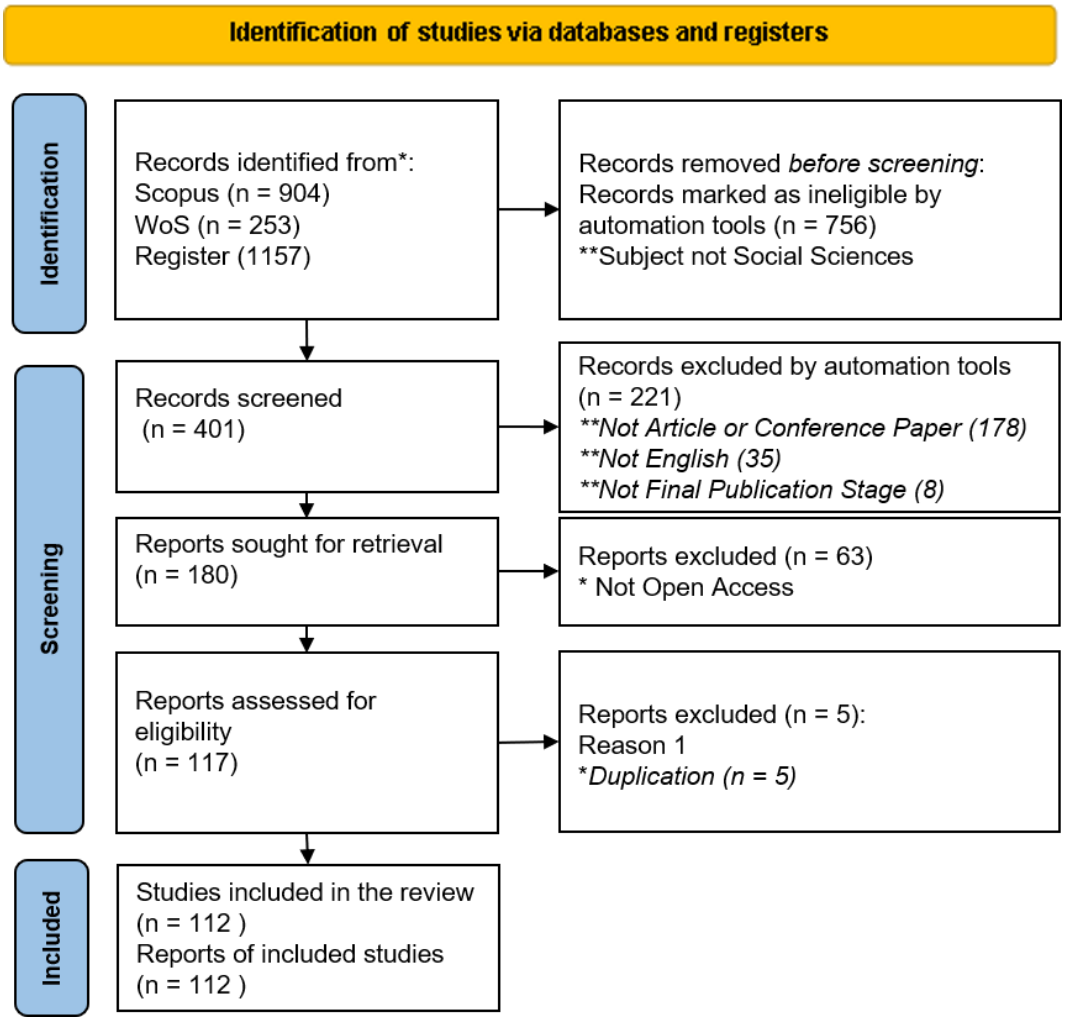



**Fig. 1. PRISMA flow diagram of study identification and selection.**

### 3. Data handling and bibliometric analysis

The cleaned merged export provides titles, publication years, authors, affiliations, citations, abstracts, and keyword fields. Original numerical indicators are retained in the analytical tables. Duplicate handling is documented in Fig. 1; additional undocumented harmonization procedures are not assumed. In particular, separately reported keywords such as generative AI and generative artificial intelligence remain distinct in the network because the exported representation treats them separately. This choice preserves the map while limiting the interpretation of apparent thematic separation.

RQ1 is addressed through total publications (TP), total citations (TC), number of cited publications (NCP), author and institutional counts, and citation indices. Citations per publication (C/P) divide TC by TP; citations per cited publication (C/CP) divide TC by NCP. These ratios describe different denominators. The h-index records the largest number of publications with at least that many citations. Additional platform-reported indices are reproduced without treating them as measures of research quality. Country and author participation may overlap because a publication can have multiple affiliations and contributors.

Annual statistics answer RQ2 by grouping publications and their accumulated citations by publication year. Country statistics address RQ3, and the ranked document table addresses RQ4. The three-field representation addresses RQ5 by examining associations among countries, selected authors, and author keywords. It is not a co-authorship network and does not demonstrate collaboration between countries. RQ6 uses author-keyword co-occurrence to identify recurring vocabulary and its connections. The contextual synthesis then considers selected primary studies and relevant reviews alongside those descriptive results, preserving the distinction between metadata associations and findings reported within individual studies.

### Data Analysis

The numerical, three-field, and network exports identify BiblioSpy (Ahmi, 2026) as the analytical platform. The accompanying network configuration specifies author keywords, full counting, association normalization, and a minimum occurrence threshold of three. It records resolution 1.0, a minimum cluster size of 5, 10 random starts, and 10 clustering iterations, with a clustering seed of 0. The map contains 26 retained terms in four clusters. Its node-level occurrence, link, and total-link-strength fields support the descriptions in the Results section.

The map and network files are distributed in a compatible exchange format; that format does not, by itself, establish that another software package was used to generate the analysis. Cluster labels in this manuscript are interpretive summaries of documented keyword memberships. They are not latent variables, validated taxonomies of all included studies, or causal pathways. Citation counts and network values refer to the export snapshot, and no longitudinal citation-accrual model or statistical significance test is applied.

## RESULTS

### 1. Corpus Profile

Table 3 addresses RQ1 and describes 112 publications with 3,035 citations. Of these publications, 103 had received at least one citation, representing 91.96% of the corpus; nine were uncited in the export snapshot. The reported citation averages were 27.1 per publication and 29.47 per cited publication. The difference follows from excluding uncited documents from the second denominator. The dataset's h-index was 25, and the citation sum within the h-core was 2,465. These measures describe citation concentration within the selected literature, rather than the effectiveness of the applications discussed.

The corpus involved 330 unique authors and 335 contributing-author occurrences across 184 institutions and 43 countries. Twenty-three documents were single-authored, while the reported average number of authors per paper was 2.99. International co-authorship accounted for 21.43% of publications. The author-occurrence and unique-author counts should remain separate because the same person can contribute to more than one record. Institution and country counts likewise summarize affiliation participation without indicating where technologies were deployed or communities were studied.

The 387 unique author keywords and 17,119 references indicate the volume of descriptive and cited material associated with the corpus. They do not establish its conceptual diversity without further coding. Table 3 also reports an annual growth rate of 71.58% and a mean document age of 2.03 years. These summary indicators are consistent with the strong recent concentration of output examined in the annual analysis, although a growth statistic does not establish the maturity of a research area.

**Table 3. Bibliometric summary.**

| Indicator | Value |
|---|---|
| Timespan | 2021:2025 |
| Total Publications (TP) | 112 |
| Citable Year Span | 5 years |
| Annual Growth Rate | 71.58% |
| Document Average Age | 2.03 years |
| Total Citations (TC) | 3035 |
| Average Citations per Document | 27.1 |
| Number of Cited Papers (NCP) | 103 |
| Average Citations per Cited Paper | 29.47 |
| Average Citations per Year | 607 |
| Total Contributing Authors | 335 |
| Total Unique Authors | 330 |
| Total Institutions | 184 |
| Total Countries | 43 |
| Single-Authored Documents | 23 |
| Average Authors per Paper | 2.99 |
| International Co-authorship Rate | 21.43% |
| Unique Author Keywords | 387 |
| Total References | 17119 |
| h-index | 25 |
| g-index | 53 |
| m-index | 4.167 |
| e-index | 42.9 |
| Citation Sum within h-Core | 2465 |

**Source: Generated by the author(s) using BiblioSpy® (Ahmi, 2026).**

## 2. Publication and Citation Patterns

Fig. 2 and Table 4 address RQ2. Annual output increased from six publications in 2021 to ten in 2022, 17 in 2023, 27 in 2024, and 52 in 2025. The final two years contributed 79 of 112 publications, or 70.54%. The largest absolute increase occurred between 2024 and 2025, when output rose by 25 publications. Within the observed period, the publication series therefore shows sustained expansion followed by a particularly large final-year increase.

The citation distribution differs from the output distribution. Publications from 2022 accumulated 920 citations, the highest annual-cohort total, followed by 633 for 2021. Citations per publication were highest for the 2021 cohort at 105.5 and declined to 7.75 for 2025. The 2025 cohort nevertheless contributed 403 citations

and included 45 cited publications. The series groups citations by the year of publication; it does not represent the number of citations received in each calendar year.

Older publications have had more time to accumulate citations, while the latest cohort contributes the most records. The contrast between rising output and lower recent citation averages is therefore compatible with unequal citation exposure. It cannot establish reduced research quality or diminished policy relevance in 2025. The annual table distinguishes between the growing volume of indexed scholarship and the distribution of its citation visibility as of the export date.

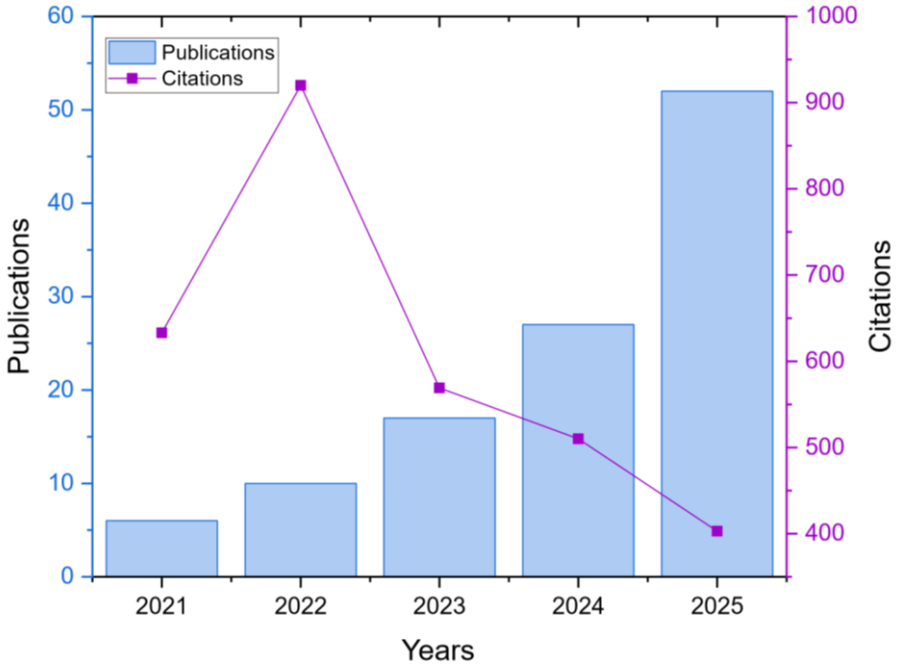



**Fig. 2. Annual publications and accumulated citations by publication year.**

**Table 4. Annual bibliometric statistics, 2021–2025.**

| Year | TP | NCA | NCP | TC | C/P | C/CP | h-index | g-index | m-index |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | 6 | 15 | 6 | 633 | 105.5 | 105.5 | 5 | 6 | 0.833 |
| 2022 | 10 | 27 | 10 | 920 | 92 | 92 | 9 | 10 | 1.8 |
| 2023 | 17 | 46 | 16 | 569 | 33.47 | 35.56 | 11 | 17 | 2.75 |
| 2024 | 27 | 88 | 26 | 510 | 18.89 | 19.62 | 12 | 22 | 4 |
| 2025 | 52 | 159 | 45 | 403 | 7.75 | 8.96 | 8 | 18 | 4 |

**Note. TP = total publications; NCA = contributing authors; NCP = cited publications; TC = total citations; C/P = citations per publication; C/CP = citations per cited publication. Source: Generated by the author(s) using BiblioSpy® (Ahmi, 2026).**

## 3. Highly Cited Documents

Table 5 addresses RQ4. Busuioc, (2021) analysis of algorithmic accountability leads the ranked documents with 518 citations. The study on AI use in government across the European Union follows with 321, and Gesk & Leyer, (2022) public-service acceptance study has 176. These leading topics concern accountability, adoption, and citizen responses. Their prominence indicates that the citation structure of the retrieved literature is closely linked to administrative and governance questions.

The remaining highly cited documents extend this orientation. They address social inclusion, organizational implementation, public administration research, AI assimilation, participatory governance, service delivery efficiency, and trustworthy decision support. Citation totals among these entries range from 133 to 82. Maragno et al., (2023), for example, contribute an implementation perspective, while Berman et al., (2024) examine a public employment-service system. These are distinct research contributions and should not be ranked as if they assessed the same intervention or used a common outcome measure.

The table also clarifies the study boundary. The ten leading titles do not collectively define a vision-only evidence base. Their relevance lies in the institutional environment in which visual technologies might be used and assessed. The ranking therefore supports interpretation of the broader AI–administration intersection retrieved by the query. Citation prominence identifies visibility within that corpus; it does not verify a technology's operational reliability, quantify a community benefit, or indicate whether a publication's conclusions have been replicated.

Table 5. Ten highly cited documents in the supplied dataset.

| No. | Author(s) | Title | Source Title | TC |
|---|---|---|---|---|
| 1 | Busuioc (2021) | Accountable Artificial Intelligence: Holding Algorithms to Account | Public Administration Review | 518 |
| 2 | Van & Misuraca (2022) | Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union | Government Information Quarterly | 321 |
| 3 | Gesk & Leyer (2022) | Artificial intelligence in public services: When and why citizens accept its usage | Government Information Quarterly | 176 |
| 4 | Djatmiko et al. (2025) | Digital Transformation and Social Inclusion in Public Services: A Qualitative Analysis of E-Government Adoption for Marginalized Communities in Sustainable Governance | Sustainability (Switzerland) | 133 |
| 5 | Maragno et al. (2023) | Exploring the factors, affordances and constraints outlining the implementation of Artificial Intelligence in public sector organizations | International Journal of Information Management | 120 |
| 6 | Mcdonald B.d. et al. (2022) | The future of public administration research: An editor's perspective | Public Administration | 118 |
| 7 | Alshahrani et al. (2022) | An attention-based view of AI assimilation in public sector organizations: The case of Saudi Arabia | Government Information Quarterly | 102 |
| 8 | Moon (2023) | Searching for inclusive artificial intelligence for social good: Participatory governance and policy recommendations for making AI more inclusive and benign for society | Public Administration Review | 96 |
| 9 | Kulal et al. (2024) | Enhancing public service delivery efficiency: Exploring the impact of AI | Journal of Open Innovation: Technology, Market, and Complexity | 91 |
| 10 | Berman et al. (2024) | Trustworthy AI in the public sector: An empirical analysis of a Swedish labor market decision-support system | Technology in Society | 82 |

*Note: TC = Total Citations*

*Source: Generated by the author(s) using BiblioSpy® (Ahmi, 2026).*

## 4. Country Contributions

Fig. 3 and Table 6 answer RQ3. Indonesia and Spain each contribute ten publications, followed by China with eight. The Netherlands and the United States each contribute seven, while India, Italy, the United Kingdom, and Romania each have six. The listed countries therefore show a relatively narrow range of publication counts, alongside much larger differences in citation totals.

The Netherlands records 737 citations and 105.29 citations per publication, the highest values among the countries listed in Table 6. Italy follows with 482 citations and an average of 80.33, while the United States records 284 citations and an average of 40.57. Indonesia and Spain have equal publication totals but different

citation profiles, at 199 and 45 citations, respectively. These comparisons show why productivity should be distinguished from average citation visibility. Small sets of highly cited papers can substantially affect country-level averages.

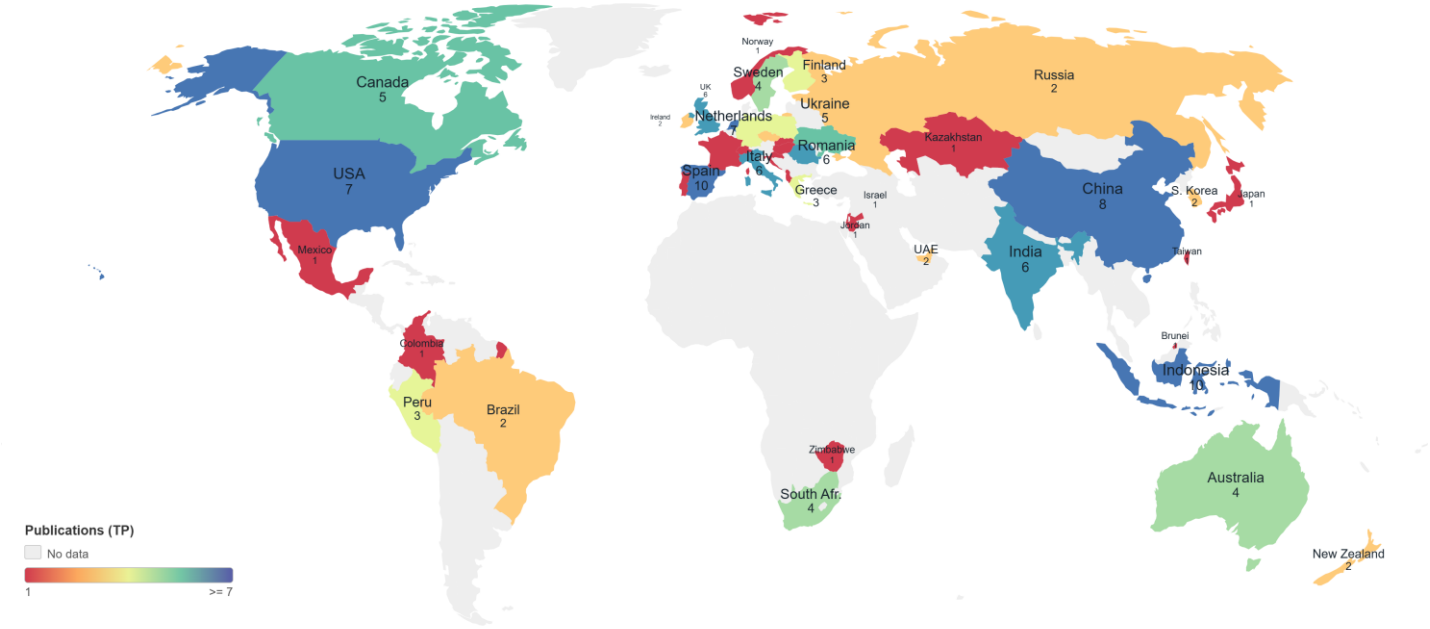



**Fig. 3. Geographic distribution of publication contributions.**

*Source: Generated by the author(s) using BiblioSpy® (Ahmi, 2026).*

**Table 6. Countries with at least six publications.**

| Country | TP | NCA | NCP | TC | C/P | C/CP | h | g | m |
|---|---|---|---|---|---|---|---|---|---|
| Indonesia | 10 | 26 | 8 | 199 | 19.9 | 24.88 | 5 | 10 | 1.25 |
| Spain | 10 | 21 | 9 | 45 | 4.5 | 5 | 5 | 6 | 1.25 |
| China | 8 | 32 | 8 | 84 | 10.5 | 10.5 | 3 | 8 | 0.5 |
| Netherlands | 7 | 10 | 7 | 737 | 105.29 | 105.29 | 5 | 7 | 0.833 |
| United States | 7 | 17 | 7 | 284 | 40.57 | 40.57 | 6 | 7 | 1.2 |
| India | 6 | 18 | 6 | 175 | 29.17 | 29.17 | 3 | 6 | 0.75 |
| Italy | 6 | 13 | 6 | 482 | 80.33 | 80.33 | 5 | 6 | 1 |
| United Kingdom | 6 | 12 | 6 | 57 | 9.5 | 9.5 | 5 | 6 | 1.25 |
| Romania | 6 | 11 | 5 | 151 | 25.17 | 30.2 | 4 | 6 | 1 |

*Note: TP = Total Publications, NCA = Number of Contributing Authors, NCP = Number of Cited Publications, TC = Total Citations, C/P = Citations per Publication, C/CP = Citations per Cited Publication, h = h-index, g = g-index, m = m-Quotient. Source: Generated by the author(s) using BiblioSpy® (Ahmi, 2026).*

## 5. Country–author–keyword relationships

Fig. 4 addresses RQ5 and links selected countries, authors, and author keywords. The three-field export associates Romanian authors with artificial intelligence, public administration, public services, digital transformation, governance, and related terms. The Netherlands is linked to authors whose selected keywords include artificial intelligence. Other displayed connections associate Spain with artificial intelligence, automation, and innovation, and Greece with large language models, public service delivery, e-government, and machine learning.

These paths illustrate how topical vocabulary is distributed across selected author-affiliation records. They do not establish that authors connected to the same term collaborated, share an institutional affiliation, or reached similar conclusions. The export also contains incomplete paths, including an author without a country value and a country–author entry without an author keyword. Such entries should be treated as incomplete field coverage. They provide no basis for assigning an unreported affiliation or inferring that the author has no substantive research topic.

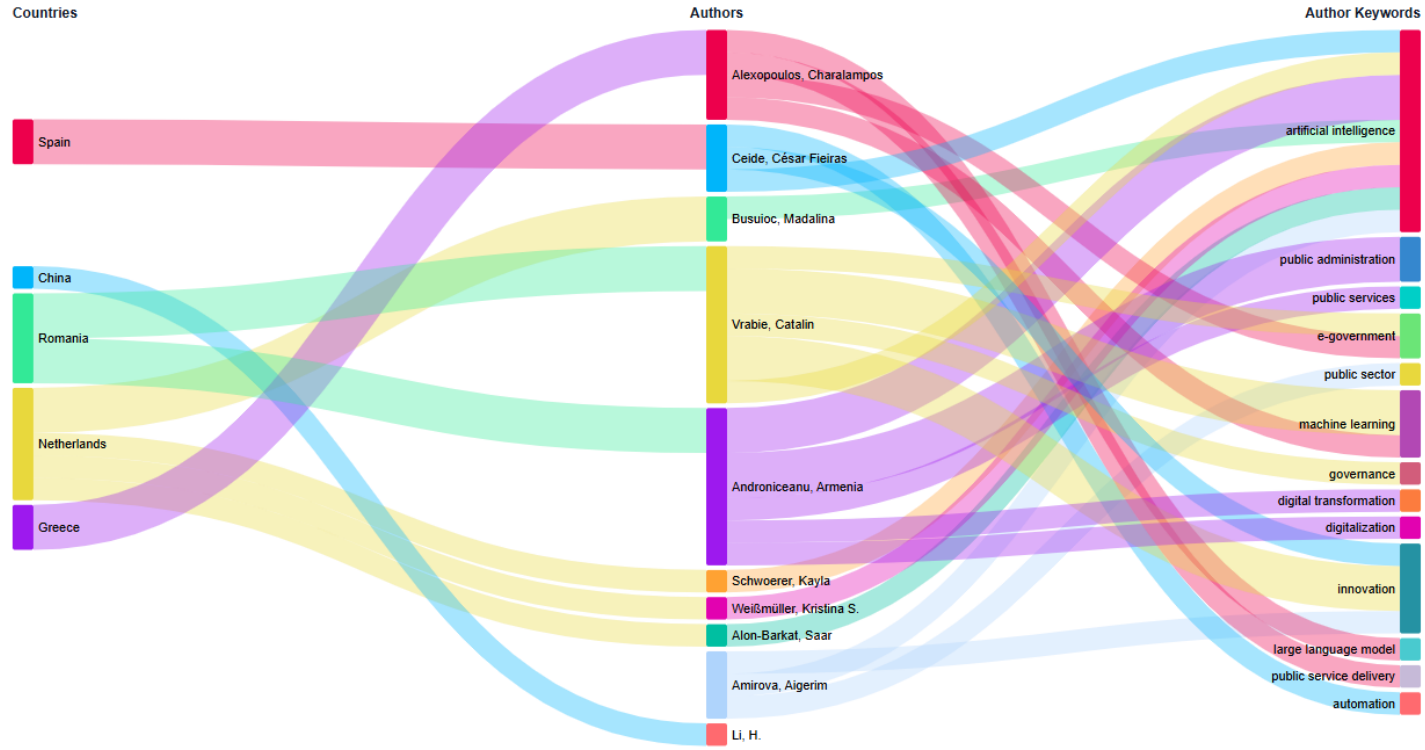



Fig. 4. Country–author–author keyword relationships.

## 6. Keyword Co-occurrence Analysis

Fig. 5 addresses RQ6 by assigning 26 keywords to four clusters. Artificial intelligence is the most frequent retained term, occurring in 51 documents, with 23 links and a total link strength of 61. Public administration occurs in 14 documents, public services in 12, and e-government in ten. These frequencies describe metadata usage rather than the prevalence of evaluated interventions.

Cluster 1, shown in magenta, groups artificial intelligence with big data, governance, innovation, machine learning, public sector, and public service delivery. Its seven terms suggest a theme of AI-enabled public-sector activity. Cluster 2, in blue, contains government, public administration, public services, technology, and transparency. Its five terms connect administrative settings with the visibility of decision processes.

Cluster 3, in green, contains eight terms: AI governance, ChatGPT, generative AI, generative artificial intelligence, large language model, public policy, smart city, and social impact. This grouping connects emerging AI applications to policy and community-facing concerns. The presence of both generative-AI formulations also shows that related labels remain separate in the export. Cluster 4, in yellow, contains automation, digital governance, digital transformation, digitalization, e-government, and the Sustainable Development Goals. It represents six terms associated with digital-government transformation.

Artificial intelligence is linked to most retained terms. These connections do not establish a progression from technical adoption to social benefit. No retained keyword explicitly names computer vision or machine vision. That absence is specific to the thresholded map and cannot establish that visual applications are absent from every record. The conceptual structure is predominantly administrative and governance-oriented, which limits a vision-exclusive interpretation.

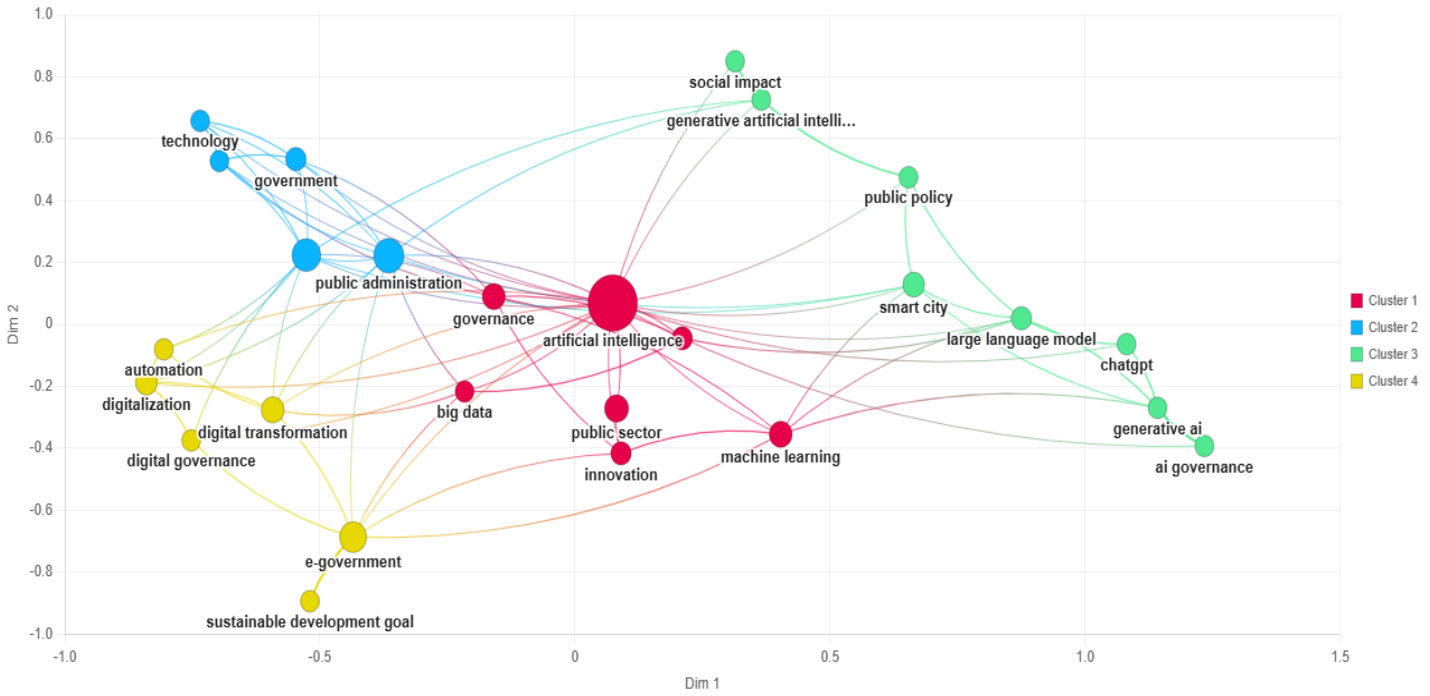



**Fig. 5. Author-keyword co-occurrence network with at least three occurrences.**

*Source: Generated by the author(s) using BiblioSpy® (Ahmi, 2026).*

# DISCUSSION

## 1. Growth of administrative-AI research and the public-sector AI research landscape

The annual results show that research activity expanded substantially toward the end of the observed period. The simultaneous prominence of accountability and public-service acceptance among highly cited documents suggests a field concerned with the institutional conditions of AI use. These patterns answer different questions: publication counts indicate increasing indexed attention, while citations identify visible contributions within that literature. Neither measures the extent to which public organizations have successfully deployed AI systems or improved service performance.

The search structure helps explain the relationship between the intended topic and the observed findings. Broad AI terms are combined with administration- and community-facing concepts, so work on predictive tools, conversational services, and organizational governance can be entered without a visual component. Caiza et al., 2024) similarly approach governmental decision-making through broad AI coverage, whereas Nemavhola et al., (2025) impose a specific face-recognition boundary. The present map corresponds to the broader public-sector AI scope. Comparisons across applications would require study-level coding of the administrative task, implementation setting, and outcome measure.

This breadth is analytically useful because it identifies the surrounding scholarship that public-service evaluation can draw on. It also limits claims that the current results establish AI's impact. Stiernströmer, (2026)

The review shows that public perceptions and operational outcomes constitute uneven parts of the evidence within a single domain-specific application. The present study therefore contributes an administrative research map and an evaluation agenda. A more detailed outcome synthesis would need to establish which records describe deployed AI systems and what each system was used to do.

## 2. AI-enabled public-sector activity: implementation before benefit

The first keyword cluster links AI and machine learning to public-sector activity, governance, and service

delivery. This association makes organizational implementation a reasonable interpretive theme, but keyword co-occurrence cannot show whether a tool became part of routine work. Maragno et al., (2023) supply a more specific account through interrelated technological, organizational, and environmental factors. Alshahrani et al., (2022) further show how managerial attention and operational attention shape assimilation. Their findings suggest that implementation should be evaluated as a process shaped by institutional conditions, rather than simply recorded as the presence of an AI system.

For an AI-enabled administrative workflow, that process includes the point at which an output enters a case or workflow, the staff member authorized to act on it, and the procedure for resolving uncertainty. These are proposed evaluation dimensions derived from the organizational literature, not observed characteristics of all 112 records. Their value lies in connecting an application to the work it changes. A study reporting improved model accuracy could then be read alongside evidence about review time, error correction, staff workload, and decisions made after human verification.

The distinction also matters when interpreting efficiency claims. Kulal et al., (2024) provide chief information officers' assessments of public-service delivery and infrastructure readiness. Such evidence is useful for understanding organizational experience, but it should be complemented by independent measures of workflow. Future administrative-AI studies could examine whether time saved at initial recognition is offset by checking uncertain matches or handling complaints. This would turn an assumed efficiency benefit into a testable administrative outcome without dismissing the informative role of practitioner perceptions.

### 3. Public administration and transparency: accountable use of automated outputs

The second cluster connects public administration and public services with transparency. The highly cited document ranking reinforces this concern through the prominence of Busuioc, (2021) accountability analysis. Together, these metadata patterns indicate sustained attention to the institutional justification of automated decisions. They do not establish that transparent procedures have been implemented or that citizens receive effective remedies when systems fail.

Berman et al., (2024) provide an empirical comparison: the Swedish employment-service case shows difficulties in making decision support sufficiently interpretable and understandable within the organization. Alon-Barkat et al., (2025) address a different level, showing that citizens continue to hold public organizations responsible for discriminatory algorithmic outcomes. The two studies should not be collapsed into a single mechanism. One concerns an implemented organizational system; the other experimentally examines judgments of responsibility. Their combined relevance is that technical delegation leaves both practical understanding and public accountability unresolved.

For AI-enabled administrative workflows, this suggests evaluating how evidence is communicated, who can contest an automated recommendation, and whether officials can recognize unsuitable uses of the output. Reporting only acceptance or confidence would leave these procedural questions unanswered. A useful research design would document the decision path from data input to final administrative action and compare error handling across user groups. The resulting evidence could distinguish a technically plausible explanation from a procedure that affected people can actually use to seek review.

### 4. Generative AI, policy, and the limits of technological transfer

The third cluster places language-model terms alongside public policy, smart city, and social impact. Its composition indicates attention to generative AI within the mapped public-sector research. Kleiman & Barbosa, (2025) chatbot case illustrates one adjacent application: early user experiences suggest potential for more efficient responses about a federal program, but further refinement is needed. That finding concerns language-based assistance in a specific administrative setting, it does not establish equivalent benefits across other AI applications.

The broader lesson concerns the location of evaluation. A tool can appear promising in an initial service

encounter while its longer-term effects on workloads, access, or erroneous guidance remain uncertain. Gesk & Leyer, (2022) show that public-service acceptance depends on the scenario, which further cautions against treating general willingness to use AI as universal approval. Visual identity checks, informational chatbots, and risk assessments impose different requirements on users and institutions, even when they appear within the same bibliometric cluster.

A future research agenda should therefore separate modalities while retaining comparable outcome categories. Researchers could compare task completion, staff intervention, correction rates, and perceived procedural fairness across specified services, with each measure interpreted in its context. The bibliometric connection between social impact and emerging AI vocabulary identifies an area of attention, independent service and community data would be required to determine what that impact is.

## 5. Digital-government transformation and inclusion

The fourth cluster connects e-government and digital transformation with the vocabulary of automation and sustainable development. These associations suggest that AI is discussed within longer-running questions about digital public services. They do not demonstrate progress toward a particular development goal. Moon, (2023) offers a relevant normative perspective by arguing for participatory governance and attention to how AI's benefits are distributed. This contribution is a framework for considering inclusion, not an empirical estimate of improved social outcomes.

The country results provide another boundary. Indonesia and Spain have the largest numbers of listed publications, while the Netherlands has the highest listed citation total. These patterns describe the geography of research participation. They do not identify which communities are well served or excluded by AI applications. Affiliation geography is especially unsuitable as a substitute for deployment context when papers are conceptual, comparative, or based on organizations in other locations.

For community-facing AI services, inclusion should be studied at the point of use. Relevant questions concern who can access the service, who is incorrectly rejected or subjected to additional checks, and whether a workable alternative remains available. Comparative research should document actual populations and administrative contexts rather than using authors' countries as proxies. Such designs would connect the digital-government theme to observable distributional consequences while preserving the distinction between a policy aspiration and an evaluated result.

## 6. Implications and Future Research

The search retrieves a broad AI–administration intersection because its first group includes general AI terms. This design supports a public-sector research map while limiting conclusions about particular AI applications. Social science, language, document type, publication stage, and access restrictions further constrained coverage. The included records span 2021–2025 within the stated 2020–2025 scope. Database-specific search execution details and an independently auditable record-by-record selection history would be needed for exact replication.

Bibliometric indicators measure publication and citation patterns rather than study quality or operational success. Citation age affects comparisons across publication years. Country affiliations do not represent deployment locations, and co-authored publications can contribute to multiple country totals. The retained keyword network depends on thresholds and naming practices; low-frequency concepts may be absent, while related labels may remain separate. Its four clusters describe relationships in metadata, not four mutually exclusive categories of evaluated services.

The contextual evidence supports interpretation but does not constitute complete outcome charting of the 112 publications. The sources cover organizational adoption, conversational systems, administrative automation, and governance arguments. Their different methods do not justify a pooled claim of improved service performance or community benefit. The selected prior-review comparison is bounded, with later reviews used only as contextual comparators. These limitations define the additional work needed to connect the mapped public-sector AI scholarship with independently assessed administrative and community outcomes.

# CONCLUSION

This review maps a growing body of AI and public administration research relevant to public service delivery. The 112 included publications span 2021–2025, with most output appearing in the final two years. Citation visibility centers on accountability, acceptance, organizational implementation, and service delivery. Four keyword clusters connect public-sector AI, administrative transparency, generative AI and policy, and digital-government transformation.

The principal contribution to public-sector AI scholarship is that the study connects patterns of research attention with organizational implementation, administrative accountability, and service evaluation. Relevant reviews and contextual studies show why AI performance, changes in administrative work, and community consequences require separate assessment. The bibliometric patterns cannot establish that adopting AI improves efficiency, fairness, or well-being.

Future work should identify the AI application and supported task in each study, document the role of human review, and follow operational and community outcomes over time. This would connect technical evaluation with administrative responsibility and the experiences of affected people. The present map provides a bounded basis for that work while making the limits of the available evidence explicit.

# REFERENCES

1. Ahmi, A. (2026). *BiblioSpy* [Computer software]. https://bibliospy.me/
2. Alamsyah, A., & Aryfiyanto, H. (2025). Leveraging Generative AI for Public Service Innovation: A Path to Smart Government in Indonesia. *Digital Government: Research and Practice*, *6*(4), 1–36. https://doi.org/10.1145/3761820
3. Alon-Barkat, S., Busuioc, M., Schwoerer, K., & Weißmüller, K. S. (2025). Algorithmic discrimination in public service provision: Understanding citizens' attribution of responsibility for human versus algorithmic discriminatory outcomes. *Journal of Public Administration Research and Theory*, *35*(4), 469–488. https://doi.org/10.1093/jopart/muaf024
4. Alshahrani, A., Dennehy, D., & Mäntymäki, M. (2022). An attention-based view of AI assimilation in public sector organizations: The case of Saudi Arabia. *Government Information Quarterly*, *39*(4), 101617. https://doi.org/10.1016/j.giq.2021.101617
5. Berman, A., De Fine Licht, K., & Carlsson, V. (2024). Trustworthy AI in the public sector: An empirical analysis of a Swedish labor market decision-support system. *Technology in Society*, *76*, 102471. https://doi.org/10.1016/j.techsoc.2024.102471
6. Busuioc, M. (2021). Accountable Artificial Intelligence: Holding Algorithms to Account. *Public Administration Review*, *81*(5), 825–836. https://doi.org/10.1111/puar.13293
7. Caiza, G., Sanguña, V., Tusa, N., Masaquiza, V., Ortiz, A., & Garcia, M. V. (2024). Navigating Governmental Choices: A Comprehensive Review of Artificial Intelligence's Impact on Decision-Making. *Informatics*, *11*(3), 64. https://doi.org/10.3390/informatics11030064
8. Gesk, T. S., & Leyer, M. (2022). Artificial intelligence in public services: When and why citizens accept its usage. *Government Information Quarterly*, *39*(3), 101704. https://doi.org/10.1016/j.giq.2022.101704
9. Kleiman, F., & Barbosa, M. M. (2025). Management and Performance Program Chatbot: A Use Case of Large Language Model in the Federal Public Sector in Brazil. *Digital Government: Research and Practice*, *6*(2), 1–11. https://doi.org/10.1145/3700141
10. Kulal, A., Rahiman, H. U., Suvarna, H., Abhishek, N., & Dinesh, S. (2024). Enhancing public service delivery efficiency: Exploring the impact of AI. *Journal of Open Innovation: Technology, Market, and Complexity*, *10*(3), 100329. https://doi.org/10.1016/j.joitmc.2024.100329
11. Maragno, G., Tangi, L., Gastaldi, L., & Benedetti, M. (2023). Exploring the factors, affordances and constraints outlining the implementation of Artificial Intelligence in public sector organizations. *International Journal of Information Management*, *73*, 102686. https://doi.org/10.1016/j.ijinfomgt.2023.102686

12. Moon, M. J. (2023). Searching for inclusive artificial intelligence for social good: Participatory governance and policy recommendations for making AI more inclusive and benign for society. *Public Administration Review*, *83*(6), 1496–1505. https://doi.org/10.1111/puar.13648
13. Nemavhola, A., Viriri, S., & Chibaya, C. (2025). A Scoping Review of Literature on Deep Learning Techniques for Face Recognition. *Human Behavior and Emerging Technologies*, *2025*(1), 5979728. https://doi.org/10.1155/hbe2/5979728
14. Ooko, G. A. (2025). Toward a Decolonial Framework for Communicating AI in African Public Service. *Emerging Media*, *3*(3), 462–473. https://doi.org/10.1177/27523543251374174
15. Parycek, P., Schmid, V., & Novak, A.-S. (2023). Artificial Intelligence (AI) and Automation in Administrative Procedures: Potentials, Limitations, and Framework Conditions. *Journal of the Knowledge Economy*, *15*(2), 8390–8415. https://doi.org/10.1007/s13132-023-01433-3
16. Rannisto, A., & Vainionpää, F. (2025). *Trust-Affording Action: Citizens' Everyday Relations With Algorithmized Public Services*. *13*. https://doi.org/https://doi.org/10.17645/si.10001
17. Sienkiewicz-Małyjurek, K., & Zyzak, B. (2025). Digital Technologies in Public Administration Networks: Systematic Literature Review and Research Avenues. *Administration & Society*, (57). https://doi.org/10.1177/00953997251369090
18. Stiernströmer, E. (2026). Facial recognition technology in law enforcement- a scoping review of existing empirical studies. *Police Practice and Research*, 1–25. https://doi.org/10.1080/15614263.2026.2627208