Technologia Journal: Jurnal Informatika
E-ISSN: 3046-9163
Vol. 3, No. 3, 2026
DOI: https://doi.org/10.62872/rftf7g69

Integrating AI-Driven Advanced Knowledge Management Systems to Mitigate Civil Servants' Risk

Andi Agus Salim[1]✉, Hasbu Naim Syaddad[2], Luki Ishwara[3]
[1,2,3]Doctoral Program in Information Systems, Indonesian Computer University, Indonesia
e-mail:
*andiagussalim@telkomuniversuty.ac.id

ABSTRACT

The rapid digitalization of public administration has positioned artificial intelligence (AI) as a strategic lever for strengthening institutional knowledge and reducing operational risk among civil servants. Yet public organizations continue to struggle with fragmented knowledge repositories, tacit knowledge loss due to workforce turnover, inconsistent decision-making, and exposure to compliance, legal, and reputational risks arising from manual and siloed information practices. This study examines how an AI-Driven Advanced Knowledge Management System (AI-AKMS) can be integrated into civil service institutions to mitigate such risks. Using a systematic literature review of twenty-five peer-reviewed sources published between 2021 and 2026, the study synthesizes evidence on AI-enabled knowledge capture, retrieval-augmented generation, predictive risk analytics, and generative AI governance in public administration. The novelty of this study lies in proposing an integrated conceptual framework that links AI-based knowledge management functions directly to specific civil-service risk categories, namely compliance risk, decision risk, knowledge-continuity risk, and reputational risk, an integration rarely addressed jointly in prior literature. Findings indicate that AI-AKMS adoption improves knowledge retrieval accuracy, accelerates policy compliance checking, and strengthens organizational resilience, provided that governance, data quality, and human oversight mechanisms are institutionalized. The study concludes with practical implications for public-sector digital transformation strategy and identifies avenues for future empirical validation.

Keywords: *artificial intelligence; knowledge management; civil servants; public administration; risk mitigation; generative AI governance.*

INTRODUCTION

Public administration worldwide is undergoing a profound transformation as governments seek to modernize service delivery, strengthen accountability, and improve internal efficiency through digital technologies. Among these technologies, artificial intelligence (AI) has emerged as one of the most

consequential, moving beyond experimental pilots into operational deployment across tax administration, healthcare regulation, licensing, and citizen services. Civil servants, who form the backbone of daily government operations, are increasingly expected to interpret complex and frequently changing regulations, respond to citizen inquiries accurately, and make decisions that carry legal, financial, and reputational consequences. In this environment, the way government institutions manage knowledge, meaning how they capture, store, retrieve, and apply the expertise embedded in regulations, precedents, and the accumulated experience of staff, has become a determinant of institutional risk rather than merely an administrative convenience.

Traditional knowledge management (KM) systems in the public sector have generally relied on static document repositories, shared drives, and standard operating procedure manuals that require active searching by employees. Such systems assume that civil servants possess sufficient time, training, and contextual understanding to locate the correct information when needed, an assumption frequently violated by high caseloads, staff turnover, and the sheer complexity of regulatory frameworks. When knowledge is fragmented across departments or lost when experienced employees retire or transfer, government agencies face several forms of avoidable risk. Compliance risk arises when officials apply outdated or incorrect procedures. Decision risk emerges when judgments are made without access to the full context of precedent or policy intent. Knowledge-continuity risk appears as institutional memory erodes with workforce attrition. Reputational risk follows from inconsistent or erroneous responses to the public, which can undermine trust in government institutions. These risks are not hypothetical; multiple studies of AI adoption in public administration have documented uneven implementation, weak governance structures, and gaps between technological capability and institutional readiness that leave public organizations vulnerable to exactly these failure modes (Madan & Ashok, 2022; Vatamanu & Tofan, 2025).

The emergence of generative AI and large language models (LLMs) has changed what is technically possible in institutional knowledge management. Retrieval-augmented generation (RAG) architectures, in particular, allow an AI system to ground its responses in an organization's own verified documents rather than relying solely on a model's general training data, thereby reducing the risk of hallucinated or outdated guidance. Recent research on legal assistance chatbots for e-governance in low-resource language settings demonstrates that carefully evaluated RAG and fine-tuned LLM pipelines can meaningfully improve the accuracy and reliability of legal and administrative information retrieval compared with general-purpose models, while also revealing that model choice, retrieval quality, and evaluation protocol substantially affect real-world usability (Tsourma et al., 2025). In parallel, scholarship on the knowledge management perspective of generative AI argues that GenAI is reshaping the entire KM lifecycle, from knowledge creation and capture through storage, retrieval, transfer, and application, positioning AI not as a peripheral tool but as an infrastructural layer of organizational cognition (Alavi, Leidner, & Mousavi, 2024). Complementary systematic reviews of AI-enabled knowledge management describe a synergistic relationship between people and technology in which AI augments, rather than replaces, human expertise, provided that organizational processes are redesigned to accommodate the new capability (Pai et al., 2022; Alsammak, Majid, Allahham,

& Alrwashdeh, 2026).

At the same time, a distinct and growing body of literature addresses the governance, risk, and ethical dimensions of introducing AI into government functions. Analyses of AI adoption and diffusion in public administration consistently report that technical feasibility outpaces institutional capacity to manage novel risks such as algorithmic opacity, data quality deficiencies, and accountability gaps (Madan & Ashok, 2022). Frameworks for classifying sources of AI risk (Steimers & Schneider, 2022), together with broader risk-management literature associated with instruments such as the AI Act (Schuett, 2022) and international scientific consensus statements on managing extreme AI risks (Bengio et al., 2023), emphasize that risk in AI-enabled systems is multidimensional, spanning technical reliability, organizational governance, and societal trust. Studies specific to public procurement of AI systems further show that public entities frequently lack AI-specific due diligence, monitoring, and transparency mechanisms, exposing them to legal and rights-related risks that would not arise from conventional software procurement (Hickok, 2024). Complementary work on accountability mechanisms for algorithmic systems in the public sector similarly stresses that transparency and human oversight are prerequisites for legitimate AI use in government, not optional enhancements (Busuioc, 2021).

Despite this expanding body of work, three gaps remain evident in the existing literature. First, most studies of AI in knowledge management address private-sector or generic organizational contexts, while research specifically examining civil-service knowledge environments, characterized by strict legal accountability, hierarchical approval structures, and public scrutiny, remains comparatively limited. Second, the AI-in-public-administration literature tends to treat "risk" as a general governance concern rather than disaggregating it into the specific operational risk categories that knowledge management interventions can directly target, such as compliance risk, decision risk, knowledge-continuity risk, and reputational risk. Third, few studies connect the technical architecture of advanced knowledge management systems, particularly RAG-based and generative AI-enabled systems, directly to civil-service risk mitigation outcomes through an integrated conceptual model; most treat technology adoption and risk management as parallel rather than intersecting streams of inquiry, a gap also noted in broad systematic reviews of AI in public governance (Zuiderwijk, Chen, & Salem, 2021) and in recent accounts of AI's transformative but organizationally underexplored role in public-sector reform (Longo, 2024).

Evidence from adjacent public-service domains reinforces the urgency of this integration. Research on workplace AI risk management identifies psychosocial hazards, including stress, cognitive overload, and role ambiguity, that can accompany poorly governed technology adoption among frontline public employees (Howard & Schulte, 2024), while sector-specific studies of public health inspectors show that AI-driven analytics can be used constructively to predict workplace hazard and burnout when designed with employee wellbeing in view (Adamopoulos, Valamontes, Tsirkas, & Dounias, 2025). In healthcare administration, decision-support and knowledge-exchange systems built on AI have been shown to improve consistency and speed of managerial decisions in hospital and public-health settings, offering a template for how similar architectures might function within civil-service knowledge environments more

broadly (Alves, Seringa, Silvestre, & Magalhães, 2024). Studies of government-to-government knowledge exchange in AI-based public healthcare systems further illustrate that shared, AI-mediated knowledge infrastructures can enhance decision-making quality across administrative units, a pattern directly relevant to civil-service contexts where knowledge silos between departments are a persistent source of inefficiency and error (Nasseef, Baabdullah, Alalwan, Lal, & Dwivedi, 2021). Taken together, this cross-domain evidence suggests that the risk-mitigating potential of AI-enabled knowledge management is not confined to any single government function but reflects a generalizable mechanism, better information availability at the point of decision reduces the probability and severity of avoidable errors, that has yet to be systematically theorized for the civil-service workforce as a distinct population.

The novelty of the present study lies precisely in addressing this integration gap. Rather than treating AI-enabled knowledge management and civil-service risk management as separate literatures, this study synthesizes both streams into a single conceptual framework that maps specific AI-Driven Advanced Knowledge Management System (AI-AKMS) functions, including intelligent document retrieval, automated policy-compliance checking, predictive risk analytics, and AI-assisted decision support, onto the four risk categories identified above. This framework is grounded in a systematic review of twenty-five sources published predominantly within the last five years, combining literature explicitly provided for this study with additional peer-reviewed and institutional sources located independently to ensure currency and breadth of evidence. By making explicit the mechanisms through which advanced knowledge management functions reduce particular categories of institutional risk, and by identifying the governance conditions, such as human oversight, data quality assurance, and transparency, that must accompany technology adoption, this study offers both a theoretical contribution to the digital government literature and a practical reference point for public-sector leaders designing AI-enabled knowledge infrastructures. The remainder of this article presents the research method used to conduct the review, followed by the results and discussion organized around the proposed integration framework, and concludes with implications for policy and directions for future empirical research. In doing so, the study aims to move the conversation beyond the general question of whether AI should be adopted in government toward the more actionable question of how specific AI-enabled knowledge management capabilities can be configured, governed, and evaluated to protect civil servants and the institutions they serve from avoidable operational harm.

METHOD

This study employs a systematic literature review (SLR) design, selected because the research question, how AI-driven advanced knowledge management functions can be mapped onto specific civil-service risk categories, requires synthesizing a heterogeneous and rapidly evolving body of evidence rather than generating new primary data. The review follows a four-stage process adapted from the PRISMA approach to identification, screening, eligibility assessment, and inclusion, illustrated in Figure 1.

In the identification stage, records were compiled from two sources. The first source consisted of a curated bibliographic set of twenty peer-reviewed articles and conference papers, provided as the primary reference base for this study and covering

AI-enabled knowledge management, public administration digitalization, and AI risk governance. The second source consisted of an independent supplementary search conducted through Google Scholar-indexed databases (Scopus, Web of Science, ScienceDirect, SpringerLink, Wiley Online Library, and arXiv) using combinations of the keywords "artificial intelligence," "knowledge management system," "public administration," "civil servant," and "risk mitigation," restricted to publications from 2021 through 2026 with an active, resolvable digital object identifier (DOI). This supplementary search yielded seven additional sources that met the currency and accessibility criteria and were not already present in the primary set.

During screening, titles and abstracts were assessed against three inclusion criteria: (a) explicit engagement with artificial intelligence applied to knowledge management, public-sector decision-making, or organizational risk; (b) publication in a peer-reviewed journal, indexed conference proceeding, or reputable institutional report; and (c) availability of a working DOI or stable publisher link enabling verification. Sources failing to satisfy any criterion, such as opinion pieces without empirical or conceptual grounding, were excluded at this stage.

At the eligibility stage, full texts or detailed abstracts of the remaining sources were read to confirm thematic relevance to one or more of four analytical categories used to structure the synthesis: (1) AI-enabled knowledge management architectures, including retrieval-augmented generation and generative AI applications; (2) AI adoption and governance in public administration; (3) taxonomies and management frameworks for AI-related risk; and (4) civil-service and public-workforce implications of AI adoption, including workplace and decision-related risk. Sources addressing at least one category with sufficient conceptual or empirical detail were retained.

The final corpus comprised twenty-five sources, eighteen from the provided bibliographic set and seven independently located, published between 2021 and 2026, spanning journals such as Government Information Quarterly, AI & Society, Administrative Sciences, and Public Administration Review, together with institutional publications from the OECD. Data extraction focused on each source's reported AI-KM functions, identified risks, governance mechanisms, and contextual findings relevant to civil-service settings. Extracted data were then thematically synthesized using a narrative synthesis approach, in which recurring constructs were grouped into the four risk categories central to the proposed integration framework: compliance risk, decision risk, knowledge-continuity risk, and reputational risk. This synthesis directly informs the structure of the Results and Discussion section that follows.

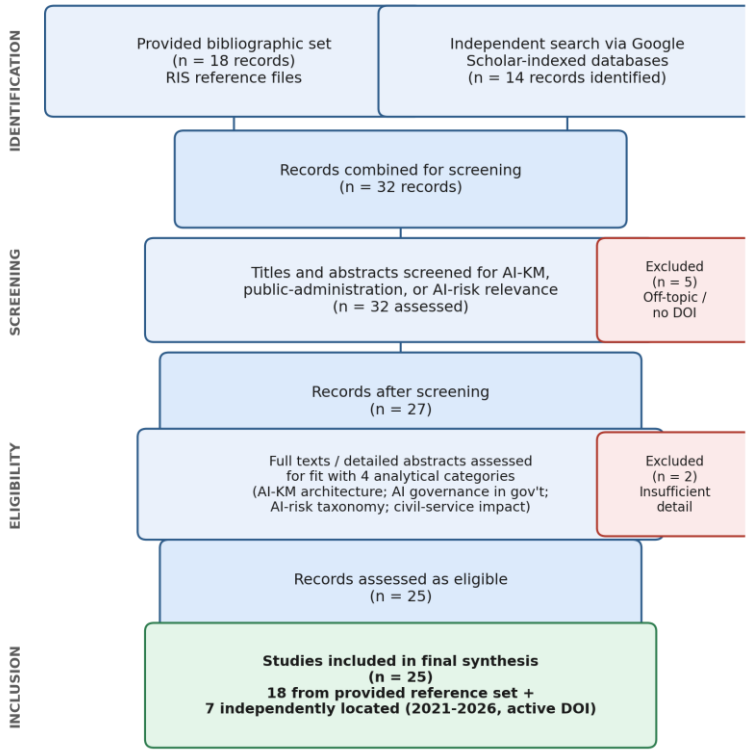

Figure 1. PRISMA-adapted flow diagram of the systematic literature review proces

RESULT AND DSCUSSION

The systematic review of twenty-five sources confirms that AI-Driven Advanced Knowledge Management Systems (AI-AKMS) intersect with civil-service risk along four recurring dimensions identified during synthesis: compliance risk, decision risk, knowledge-continuity risk, and reputational risk. Table 1 summarizes the mapping between core AI-AKMS functions, representative supporting literature, and the specific risk category each function is reported to mitigate. The discussion that follows elaborates on each category in turn before addressing the governance conditions necessary for these benefits to materialize in practice.

Table 1. Mapping of AI-AKMS Functions to Civil-Service Risk Categories

AI-AKMS Function	**Representative Sources**	**Risk Category Mitigated**	**Reported Mechanism**
Retrieval-augmented generation (RAG) for regulatory and legal queries	Tsourma et al. (2025); Alavi, Leidner, & Mousavi (2024)	Compliance risk	Grounds responses in verified regulatory text, reducing outdated or fabricated guidance
Automated policy-compliance checking and document classification	Vatamanu & Tofan (2025); Parycek, Schmid, & Novak (2023)	Compliance risk	Flags procedural deviations before administrative action is finalized
AI-assisted decision support and G2G knowledge exchange	Nasseef et al. (2021); Alves et al. (2024)	Decision risk	Provides decision-makers with synthesized precedent and cross-unit context
Predictive risk analytics for workforce and operational hazards	Adamopoulos et al. (2025); Howard & Schulte (2024)	Decision risk / workplace risk	Anticipates burnout, stress, and hazard patterns before they

			affect performance
Institutional knowledge capture and tacit-knowledge codification	Pai et al. (2022); Alsammak et al. (2026); Bratianu & Bejinaru (2025)	Knowledge-continuity risk	Preserves expertise from experienced staff against attrition and turnover
AI-KM standards integration (e.g., ISO-aligned KM and AI systems)	Khazieva, Pauliková, & Chovanová (2024)	Knowledge-continuity risk	Embeds codified knowledge into auditable, standardized organizational processes
Governance, transparency, and accountability mechanisms	Schuett (2022); Bengio et al. (2023); Busuioc (2021); Hickok (2024)	Reputational risk	Ensures oversight and explainability, sustaining public trust in AI-assisted decisions
Change-management and organizational transformation practices	Mainardi (2024); Tangi, Rodriguez Müller, & Janssen (2025); Cao et al. (2021)	Reputational risk / adoption risk	Aligns staff behavior and culture with responsible AI use, reducing implementation failure

The mapping presented in Table 1 demonstrates the main contribution of this review by connecting specific AI-AKMS functions with distinct categories of civil-service risk. Rather than assuming that AI adoption produces uniform benefits, the framework shows that different functions address different forms of institutional vulnerability. Retrieval-augmented generation and automated compliance checking primarily support compliance risk mitigation, while AI-assisted decision support and predictive analytics address decision and workplace-related risks. Institutional knowledge capture strengthens knowledge continuity, whereas transparency, accountability, and change-management mechanisms are particularly important for reducing reputational and adoption risks.

This integrated mapping also addresses an important gap in the existing literature, which has generally examined AI-enabled knowledge management, public-sector AI adoption, and organizational risk as separate areas of inquiry. The findings suggest that the effectiveness of AI-AKMS depends not only on the technical capability of the system but also on the governance conditions surrounding its use. Data quality, human oversight, and organizational readiness therefore function as cross-cutting requirements that determine whether AI-based knowledge management can genuinely reduce civil-service risk. The proposed framework consequently provides a more practical basis for public institutions to identify which AI functions should be prioritized according to specific risk categories and organizational needs.

Compliance Risk Mitigation.

The clearest evidence of risk reduction concerns compliance. Systems built on retrieval-augmented generation ground model outputs in an organization's verified regulatory corpus rather than in a model's general, and potentially outdated, training data. Evaluation of a legal-assistance chatbot for Greek e-governance found that carefully tuned RAG pipelines improved answer accuracy over unguided LLM use in low-resource-language legal contexts, though the same study cautioned that retrieval quality and evaluation protocol strongly influence real-world reliability, meaning that compliance gains are conditional on system design rather than automatic (Tsourma et al., 2025). This nuance matters for civil-service deployment: an AI-AKMS that merely wraps a general-purpose model without curated, current regulatory grounding may replicate rather than resolve compliance risk. Complementary literature on automation in administrative procedures similarly finds that AI can accelerate procedural consistency checks, but only where the underlying rule base is actively maintained and

validated against current law, underscoring that compliance-risk mitigation is a function of knowledge governance as much as of algorithmic capability (Parycek, Schmid, & Novak, 2023; Vatamanu & Tofan, 2025).

Decision Risk Mitigation.

A second consistent theme is the reduction of decision risk through improved access to synthesized precedent and cross-departmental context. Evidence from AI-based public healthcare knowledge-exchange systems demonstrates that government-to-government sharing of AI-processed information can enhance the quality and speed of managerial decisions by reducing the information asymmetry between units that traditionally operate in isolation (Nasseef, Baabdullah, Alalwan, Lal, & Dwivedi, 2021). Similarly, studies of AI-supported decision-making in hospital management show measurable improvements in decision consistency when clinicians and administrators are given AI-synthesized options rather than raw, unstructured data (Alves, Seringa, Silvestre, & Magalhães, 2024). Applied to civil-service contexts, these findings suggest that AI-AKMS functions which surface relevant precedent, cross-reference prior case outcomes, and highlight applicable policy provisions at the point of decision can reduce the incidence of inconsistent or poorly informed administrative rulings.

At the same time, predictive analytics applied to workforce hazard and burnout among public-sector inspectors illustrates a second, less obvious dimension of decision risk: decisions about staffing, workload allocation, and scheduling can themselves be improved when AI identifies emerging stress or hazard patterns, allowing managers to intervene before performance-affecting fatigue compromises judgment quality (Adamopoulos, Valamontes, Tsirkas, & Dounias, 2025). Reviews of managers' attitudes toward AI-supported organizational decision-making, however, note that trust and behavioral intention toward AI recommendations remain uneven, implying that decision-risk reduction is only partially realized unless staff are trained to interpret and appropriately weight AI-generated recommendations rather than deferring to them uncritically or dismissing them outright (Cao, Duan, Edwards, & Dwivedi, 2021).

Knowledge-Continuity Risk Mitigation.

The review provides strong support for the proposition that AI-AKMS functions can directly address knowledge-continuity risk, the erosion of institutional memory as experienced personnel exit the workforce. Systematic evidence on integrating AI into knowledge management systems highlights a synergistic relationship between technology and personnel, in which AI captures and structures tacit knowledge that would otherwise remain undocumented, then makes it retrievable independent of any single employee's presence (Pai et al., 2022). Complementary work on the nexus between AI-driven capabilities and knowledge systems in digital business environments similarly finds that organizations embedding AI within their knowledge architecture experience measurable gains in knowledge transfer efficiency (Alsammak, Majid, Allahham, & Alrwashdeh, 2026; Bratianu & Bejinaru, 2025). For civil-service institutions specifically, where succession planning is often constrained by budget and hiring cycles, this capability is particularly consequential: codifying regulatory interpretation, procedural exceptions, and case-specific reasoning into an AI-searchable knowledge base reduces dependence on any individual's tenure. Research on integrating standards for knowledge management and AI systems further indicates that pairing AI-KM

capability with recognized management-system standards improves auditability and long-term maintainability of the codified knowledge base, addressing a common failure mode in which early-stage knowledge-capture initiatives degrade over time due to inconsistent updating (Khazieva, Pauliková, & Chovanová, 2024).

Reputational Risk Mitigation And Its Governance Conditions.

The final and most conditional category concerns reputational risk, harm to public trust arising from inconsistent, opaque, or erroneous AI-assisted administrative action. Here the literature is more cautionary than celebratory. Analyses of accountability in algorithmic government systems argue that holding AI-assisted decisions to account requires deliberate transparency mechanisms, since the complexity of underlying models can otherwise obscure the basis for administrative action from both officials and the citizens affected by it (Busuioc, 2021). Studies of AI procurement in the public sector reinforce this concern, finding that public entities frequently lack AI-specific due diligence and monitoring processes, creating latent reputational exposure that may not surface until a high-profile failure occurs (Hickok, 2024). Broader scientific consensus statements on managing extreme and near-term AI risks, together with risk-management analysis of instruments such as the AI Act, converge on the view that technical reliability alone is insufficient; institutional processes for oversight, incident reporting, and human review must be embedded alongside the technology itself (Bengio et al., 2023; Schuett, 2022). Evidence from large-scale healthcare AI deployments corroborates this point outside the strict public-administration context, showing that governance structures, rather than model performance alone, determine whether enterprise-scale AI systems can be scaled safely across an organization (Bodnari & Travis, 2025). Consistent with this, research on change management for AI adoption in public administrations finds that reputational and implementation risk is substantially reduced when organizations invest deliberately in staff training, communication, and cultural adaptation rather than treating AI deployment as a purely technical rollout (Mainardi, 2024). Recent theorization of AI-augmented government transformation reinforces this point, arguing that AI integration in public administration is fundamentally a socio-technical process in which organizational structure, governance, and culture must co-evolve with the technology for benefits to be realized without introducing new risks (Tangi, Rodriguez Müller, & Janssen, 2025). Similarly, evidence on managers' behavioral intentions toward AI-supported decision-making shows that perceived trustworthiness and organizational support are stronger predictors of successful adoption than technical performance alone, reinforcing that reputational-risk mitigation is inseparable from the human and governance dimensions of implementation (Cao, Duan, Edwards, & Dwivedi, 2021).

Comparative pattern across risk categories.

Examining the four risk categories side by side reveals an important asymmetry in how directly each can be addressed by AI-AKMS functions. Compliance and knowledge-continuity risks respond relatively directly to technical intervention: grounding retrieval in verified regulatory text, or codifying tacit expertise into a searchable knowledge base, produces a fairly linear reduction in the probability that an official acts on outdated or missing information. Decision risk responds more conditionally, since the same predictive or decision-support tool can either improve or

degrade judgment quality depending on whether staff are trained to interpret its outputs critically rather than treating them as authoritative by default, a nuance highlighted repeatedly in studies of managerial attitudes toward AI-supported decision-making. Reputational risk is the most conditional of all, because it depends less on the AI system's internal accuracy than on the surrounding governance architecture, transparency to affected citizens, avenues for appeal, and mechanisms for detecting and correcting systematic errors before they accumulate into a public failure. This asymmetry has a practical implication for public-sector technology strategy: agencies seeking rapid, measurable risk reduction may reasonably prioritize compliance-oriented and knowledge-continuity-oriented AI-AKMS functions first, since these yield more predictable returns, while treating decision-support and reputational-risk-relevant functions as longer-horizon investments requiring parallel governance development rather than treating all four categories as equally ready for immediate deployment.

The Role Of Organizational Readiness

Several of the reviewed sources emphasize that the degree of risk reduction achieved is not solely a property of the technology but is substantially mediated by organizational readiness. Systematic reviews of AI adoption and diffusion in public administration document wide variance in institutional capacity, noting that agencies with mature digital infrastructure and clearer governance mandates realize more consistent benefits than those adopting AI in an ad hoc manner (Madan & Ashok, 2022; Longo, 2024). This variance helps explain why case-level findings in the literature range from strongly positive, as in specific healthcare decision-support and government-to-government knowledge-exchange implementations, to more cautionary, as in analyses of AI procurement practices that reveal persistent due-diligence gaps. For civil-service institutions considering AI-AKMS adoption, this suggests that a readiness assessment, covering data infrastructure, staff digital literacy, and existing governance mechanisms, should precede large-scale deployment rather than following it, since retrofitting governance onto an already-deployed system is considerably more difficult than embedding it from the outset. Frameworks proposed for classifying sources of AI risk provide a useful diagnostic starting point for such assessments, distinguishing risks originating in data, in the model itself, in the interaction between system and user, and in the broader organizational and societal context (Steimers & Schneider, 2022), a categorization that maps reasonably well onto the four civil-service risk categories synthesized in this study. Broader analyses of trustworthy AI requirements likewise emphasize that reliability, safety, and accountability must be engineered jointly rather than treated as separable properties of a system (Alzubaidi et al., 2023).

Limitations Of The Current Evidence Base.

It is important to note that much of the reviewed literature remains conceptual, exploratory, or based on qualitative expert interviews rather than large-sample quantitative evaluation of civil-service outcomes specifically. Several of the most directly relevant studies, including the theorization of AI-augmented government transformation and the systematic review of AI-KM synergy, explicitly call for further empirical validation across diverse public-sector contexts. Similarly, evidence on RAG-based legal assistance systems, while methodologically rigorous, is drawn from a single national language setting and administrative domain, limiting immediate

generalizability to other civil-service environments with different legal traditions, languages, and organizational structures. These limitations do not undermine the conceptual mapping proposed in this study but do indicate that the specific magnitude of risk reduction reported here should be interpreted as directionally supported by current evidence rather than precisely quantified, pending future empirical studies designed explicitly around civil-service populations.

Cross-Cutting Implications.

Read together, the findings across all four categories reveal a consistent pattern: AI-AKMS functions can reduce specific, identifiable categories of civil-service risk, but the magnitude of risk reduction is conditional on three cross-cutting governance requirements evident across the reviewed literature. First, data quality and currency must be actively maintained, since AI systems grounded in outdated or incomplete regulatory corpora can introduce new compliance risk even as they appear to reduce it. Second, human oversight must remain embedded in the decision workflow, particularly for decisions with legal or rights-related consequences, consistent with broader AI risk-management guidance emphasizing human-in-the-loop design for high-stakes applications. Third, organizational change management, encompassing training, communication, and incentive alignment, determines whether the codified knowledge and predictive capabilities generated by AI-AKMS are actually adopted into daily practice or remain underused technical assets. These three conditions collectively explain why the empirical and conceptual literature reviewed here reports mixed rather than uniformly positive outcomes for AI adoption in public administration: the technology's risk-mitigating potential is real, but it is realized only where institutions treat knowledge governance, human oversight, and organizational change as integral components of system design rather than as afterthoughts to a primarily technical implementation

CONCLUSION

This study set out to examine how AI-Driven Advanced Knowledge Management Systems can be integrated into civil-service institutions to mitigate operational risk, and to address a gap in the literature where AI-enabled knowledge management and public-sector risk management have largely been studied as separate concerns. Through a systematic review of twenty-five sources published predominantly between 2021 and 2026, this study proposed and substantiated an integrated conceptual framework linking specific AI-AKMS functions, retrieval-augmented generation, automated compliance checking, AI-assisted decision support, predictive workforce analytics, and tacit-knowledge codification, to four civil-service risk categories: compliance risk, decision risk, knowledge-continuity risk, and reputational risk. The evidence indicates that these functions can meaningfully reduce risk exposure, but only where three governance conditions are institutionalized alongside the technology: sustained data quality and currency, embedded human oversight for consequential decisions, and deliberate organizational change management encompassing staff training and cultural adaptation. For public-sector leaders, the practical implication is that AI-AKMS should be conceived not as a standalone technical upgrade but as a socio-technical program in which knowledge governance and human capability development proceed in parallel with system

deployment. For researchers, the study points toward a clear agenda for future empirical work, including longitudinal studies measuring actual compliance error rates, decision consistency, and institutional-memory retention before and after AI-AKMS adoption across diverse civil-service contexts, cross-national comparative studies examining how differing legal and administrative traditions shape the risk-mitigation pathways identified here, and mixed-methods research capturing civil servants' own perceptions of trust, usability, and behavioral change when working alongside AI-enabled knowledge systems. As governments continue to expand AI adoption in core administrative functions, grounding this expansion in an explicit understanding of which risks are mitigated, and under what governance conditions, offers a more defensible basis for public-sector digital transformation than adoption driven primarily by technological availability

powerful tools are designed, integrated, and governed in ways that are technically sound, organizationally effective, and aligned with broader societal values

REFERENCES

Adamopoulos, I., Valamontes, A., Tsirkas, P., & Dounias, G. (2025). Predicting workplace hazard, stress and burnout among public health inspectors: An AI-driven analysis in the context of climate change. European Journal of Investigation in Health, Psychology and Education, 15. https://doi.org/10.3390/ejihpe15050065

Alavi, M., Leidner, D., & Mousavi, R. (2024). Knowledge management perspective of generative artificial intelligence. Journal of the Association for Information Systems, 25, 15. https://doi.org/10.17705/1jais.00859

Alsammak, M., Majid, A. S. A., Allahham, M., & Alrwashdeh, M. (2026). The nexus between AI-driven capabilities and knowledge systems in digital business environments. Knowledge and Performance Management, 10(1). https://doi.org/10.21511/kpm.10(1).2026.09

Alves, M., Seringa, J., Silvestre, T., & Magalhães, T. (2024). Use of artificial intelligence tools in supporting decision-making in hospital management. BMC Health Services Research, 24. https://doi.org/10.1186/s12913-024-11602-y

Alzubaidi, L., Al-Sabaawi, A., Bai, J., Dukhan, A., Alkenani, A., Al-Asadi, A., Alwzwazy, H., Manoufali, M., Fadhel, M., Albahri, A., Moreira, C., Ouyang, C., Zhang, J., Santamaría, J., Salhi, A., Hollman, F., Gupta, A., Duan, Y., Rabczuk, T., Abbosh, A., & Gu, Y. (2023). Towards risk-free trustworthy artificial intelligence: Significance and requirements. International Journal of Intelligent Systems, 2023, 1–41. https://doi.org/10.1155/2023/4459198

Bengio, Y., Hinton, G., Yao, A., Song, D., Abbeel, P., Darrell, T., Harari, Y., Zhang, Y.-Q., Xue, L., Shalev-Shwartz, S., Hadfield, G., Clune, J., Maharaj, T., Hutter, F., Baydin, A., McIlraith, S., Gao, Q., Acharya, A., Krueger, D., Dragan, A., Torr, P., Russell, S., Kahneman, D., Brauner, J., & Mindermann, S. (2023). Managing extreme AI risks amid rapid progress. Science, 384(6698), 842–845. https://doi.org/10.1126/science.adn0117

Bodnari, A., & Travis, J. (2025). Scaling enterprise AI in healthcare: The role of governance in risk mitigation frameworks. NPJ Digital Medicine, 8. https://doi.org/10.1038/s41746-025-01700-4

Bratianu, C., & Bejinaru, R. (2025). Integrating artificial intelligence and artificial knowledge in the knowledge management systems. Proceedings of the International Conference on Business Excellence, 19(1), 3171–3179. https://doi.org/10.2478/picbe-2025-0242

Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. Public Administration Review, 81(5), 825–836. https://doi.org/10.1111/puar.13293

Cao, G., Duan, Y., Edwards, J., & Dwivedi, Y. K. (2021). Understanding managers' attitudes and behavioral intentions towards using artificial intelligence for organizational decision-making. Technovation, 106, 102312. https://doi.org/10.1016/j.technovation.2021.102312

Hickok, M. (2024). Public procurement of artificial intelligence systems: New risks and future proofing. AI & Society, 39(3), 1213–1227. https://doi.org/10.1007/s00146-022-01572-2

Howard, J., & Schulte, P. (2024). Managing workplace AI risks and the future of work. American Journal of Industrial Medicine. Advance online publication. https://doi.org/10.1002/ajim.23653
https://doi.org/10.1109/smartindustrycon65166.2025.10986153

Khazieva, N., Paulíková, A., & Chovanová, H. (2024). Maximising synergy: The benefits of a joint implementation of knowledge management and artificial intelligence system standards. Machine Learning and Knowledge Extraction, 6(4), 2282–2302. https://doi.org/10.3390/make6040112

Longo, J. (2024). The transformative potential of artificial intelligence for public sector reform. Canadian Public Administration, 67(4), 1–20. https://doi.org/10.1111/capa.12587

Madan, R., & Ashok, M. (2022). 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

Mainardi, I. (2024). Change management: Artificial intelligence (AI) at the service of public administrations. AI & Society, 40, 3953–3981. https://doi.org/10.1007/s00146-024-02136-2

Nasseef, O., Baabdullah, A. M., Alalwan, A. A., Lal, B., & Dwivedi, Y. K. (2021). Artificial intelligence-based public healthcare systems: G2G knowledge-based exchange to enhance the decision-making process. Government Information Quarterly, 39(4), 101618. https://doi.org/10.1016/j.giq.2021.101618

Organisation for Economic Co-operation and Development. (2025). Governing with artificial intelligence: The state of play and way forward in core government functions. OECD Publishing. https://doi.org/10.1787/795de142-en

Pai, R., Shetty, A., Shetty, A., Bhandary, R., Shetty, J., Nayak, S., Dinesh, T. K., & D'souza, K. J. (2022). Integrating artificial intelligence for knowledge management systems – synergy among people and technology: A systematic review of the evidence. Economic Research-Ekonomska Istraživanja, 35(1), 7043–7065. https://doi.org/10.1080/1331677x.2022.2058976

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, 1–26. https://doi.org/10.1007/s13132-023-01433-3

Schuett, J. (2023). Risk management in the Artificial Intelligence Act. European Journal of Risk Regulation, 1–19. https://doi.org/10.1017/err.2023.1

Segovia, M., & García, J. (2022). Design, modeling and implementation of digital twins. Sensors, 22. https://doi.org/10.3390/s22145396

Semeraro, C., Lezoche, M., Panetto, H., & Dassisti, M. (2021). Digital twin paradigm: A systematic literature review. Computers in Industry, 130, 103469. https://doi.org/10.1016/j.compind.2021.103469

Štefko, R., Frajtova-Michalikova, K., Straková, J., & Novák, A. (2025). Digital twin-based virtual factory and cyber-physical production systems, collaborative autonomous robotic and networked manufacturing technologies, and enterprise and business intelligence algorithms for industrial metaverse. Equilibrium. Quarterly Journal of Economics and Economic Policy. https://doi.org/10.24136/eq.3557

Steimers, A., & Schneider, M. (2022). Sources of risk of AI systems. International Journal of Environmental Research and Public Health, 19(6), 3641. https://doi.org/10.3390/ijerph19063641

Sueldo, C. S., Villar, S., De Paula, M., & Acosta, G. (2021). Integration of ROS and Tecnomatix for the development of digital twins based decision-making systems for smart factories. IEEE Latin America Transactions, 19, 1546–1555. https://doi.org/10.1109/tla.2021.9468608

Tangi, L., Rodriguez Müller, A. P., & Janssen, M. (2025). AI-augmented government transformation: Organisational transformation and the sociotechnical implications of artificial intelligence in public administrations. Government Information Quarterly, 42(3), 102055. https://doi.org/10.1016/j.giq.2025.102055

Tsourma, M., Michail, D., Varlamis, I., Drosou, A., & Tzovaras, D. (2025). Legal assistance in low-resource languages: Evaluating RAG and fine-tuned LLMs for Greek e-governance. In 2025 3rd International Conference on Foundation and Large Language Models (FLLM) (pp. 366–373). IEEE. https://doi.org/10.1109/fllm67465.2025.11391043

Vallée, A. (2023). Digital twin for healthcare systems. Frontiers in Digital Health,

Vatamanu, A.-F., & Tofan, M. (2025). Integrating artificial intelligence into public administration: Challenges and vulnerabilities. Administrative Sciences, 15(4), 149. https://doi.org/10.3390/admsci15040149

Walton, R., Ciarallo, F., & Champagne, L. (2024). A unified digital twin approach incorporating virtual, physical, and prescriptive analytical components to support adaptive real-time decision-making. Computers & Industrial Engineering, 193, 110241. https://doi.org/10.1016/j.cie.2024.110241

Xu, J., Çelik, N., & Chen, C.-H. (2022). Real-time digital twin-based optimization with predictive simulation learning. Journal of Simulation, 18, 47–64. https://doi.org/10.1080/17477778.2022.2046520

Yan, M., Hong, L.-Y., & Warren, K. (2021). Integrated knowledge visualization and the enterprise digital twin system for supporting strategic management decision. Management Decision. https://doi.org/10.1108/md-02-2021-0182

Zenkovich, M. (2025). Digital twin-driven decision-making support approach in adaptive control of production systems. In 2025 International Russian Smart Industry Conference (SmartIndustryCon) (pp. 1130–1135). IEEE.

Zhang, N., Bahsoon, R., Tziritas, N., & Theodoropoulos, G. (2022). Explainable human-in-the-loop dynamic data-driven digital twins. In Proceedings of the International Symposium on Leveraging Applications of Formal Methods (pp. 233–243). Springer. https://doi.org/10.1007/978-3-031-52670-1_23

Zuiderwijk, A., Chen, Y.-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research

agenda. Government Information Quarterly, 38(3), 101577. https://doi.org/10.1016/j.giq.2021.101577.https://doi.org/10.1080/00207543.2021.1898691