ISSN 2409-2665
Journal of Logistics, Informatics and Service Science
Vol. 12 (2025) No. 4, pp. 129-145
DOI:10.33168/JLISS.2025.0408

Artificial Intelligence and Large Language Models in Government Document Management: A Systematic Review of Applications, Challenges, and Implementation Strategies

Huihui Yang [1,2], Haslinda Sutan Ahmad Nawi [2] and Yunfei Zhang [3, *]

[1] Nanchang Institute of Science and Technology, No. 998 Shizishan Avenue,Honggutan District, Nanchang City, Jiangxi Province, 330108, China

2 Management and Science University,University Drive, Off Persiaran Olahraga, Section 13, 40100, Shah Alam, Selangor Darul Ehsan, Malaysia

3 Innovation College, North Chiang Mai University, 169 M.3 Nongkeaw Hangdong Chiang Mai 50230

574264160@qq.com,_haslindasan@msu.edu.my,_g666501006@northcm.ac.th

Abstract. Government agencies face escalating challenges in managing vast amounts of documents, including policy papers, legal briefs, and administrative records. Traditional document management systems—reliant on manual classification and keyword-based retrieval—struggle with scalability, accuracy, and security. This study investigates how artificial intelligence (AI), extensive language models (LLMs), can transform government document management by automating classification, enhancing retrieval accuracy, and ensuring compliance with legal standards. A systematic literature review was conducted following the PRISMA guidelines. We searched Google Scholar, IEEE Xplore, ACM Digital Library, and Web of Science for peer-reviewed studies published between 2018 and 2024. Inclusion criteria focused on empirical studies, case analyses, or theoretical frameworks involving LLM or AI applications in government document management. After screening 312 records, 42 studies met the eligibility criteria and were analyzed thematically. Findings reveal that LLMs significantly outperform traditional methods in document classification (accuracy ↑ 15－30%) and natural language querying (user satisfaction ↑ 40%). Key applications include automated policy analysis, FOIA request processing, and cross-departmental knowledge retrieval. However, implementation barriers persist: technical complexity (52% of cases), staff resistance (38%), and regulatory compliance issues (29%). LLMs offer transformative potential for government document management, but success hinges on domain-specific training, robust governance frameworks, and stakeholder engagement. This study provides a practical roadmap for public sector agencies to adopt LLMs responsibly, emphasizing phased deployment, ethics oversight, and interoperability with legacy systems.

Keywords: Large Language Models, Government Document Management, Artificial Intelligence, Systematic Review, Digital Transformation, Public Sector Innovation

1. Introduction

Government agencies generate and process vast quantities of documents every day, including legal contracts, policy papers, internal reports, correspondence, and service records. As these documents' volume, complexity, and sensitivity continue to grow, traditional document management approaches—relying heavily on manual processes, physical storage, and keyword-based retrieval—are proving inadequate. These legacy systems often lead to inefficiencies, delays, errors, and security vulnerabilities, undermining administrative performance and public service delivery.

Recent digital transformation initiatives in the public sector have emphasized the need for intelligent, automated, and secure document management. Artificial Intelligence (AI) has emerged as a critical enabler in this context. In particular, Intelligent Document Management Systems (IDMS) powered by Natural Language Processing (NLP), Machine Learning (ML), and, most notably, Large Language Models (LLMs), have shown promising results in automating document classification, enabling semantic search, supporting knowledge extraction, and improving overall information governance.

Despite the rising interest in AI-driven document systems, applying LLMs in government contexts remains underexplored in the academic literature. Existing studies are often fragmented across domains, lack systematic evaluation, or focus primarily on commercial or healthcare settings. Moreover, government agencies face unique technical, legal, and organizational constraints—such as data privacy regulations, bureaucratic workflows, and legacy infrastructure—that challenge the direct adoption of private-sector AI solutions.

This study aims to address these gaps by systematically reviewing LLM-enabled document management systems within the public sector. Specifically, we examine how LLMs are applied to improve classification accuracy, enhance search and retrieval, and support digital transformation. We also identify the primary implementation barriers, including integration complexity, staff resistance, and regulatory compliance. In doing so, we provide a critical synthesis of the current state of research and practice, offering a roadmap for governments seeking to implement LLM technologies responsibly.

Research Questions

To address the gaps in current literature and practice, this study aims to answer the following research questions:

RQ1: How effective are Large Language Models (LLMs) in improving document classification accuracy compared to traditional rule-based methods in government agencies?

RQ2: What are the primary barriers (technical, organizational, legal) to implementing LLM-based document management systems in the public sector?

RQ3: What are the best practices and success factors for integrating LLMs into government document workflows?

Research Objectives

This paper aims to:

1. Systematically review the current state of AI and LLM applications in government document management.
2. Identify and analyze the key challenges and limitations reported in real-world implementations.
3. Propose practical recommendations for public sector organizations considering LLM adoption.

Contributions

This study contributes to the literature in three ways:

1. It provides a focused and up-to-date systematic review of LLM applications in government document management systems;
2. It synthesizes empirical evidence on implementation challenges and success factors from cross-sectoral case studies;
3. It offers a practical roadmap for public agencies to adopt LLMs responsibly, focusing on scalability, ethics, and policy alignment.

2. Research Method

To explore how Large Language Models (LLMs) are applied in government document management systems, a systematic literature review (SLR) was conducted per the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The review aimed to synthesize peer-reviewed studies that examine the use, benefits, challenges, and outcomes of LLMs in public sector document workflows.

2.1.Search Strategy

The literature search was conducted across four major databases: Scopus, Web of Science, IEEE Xplore, and Google Scholar. The search was limited to articles published between January 2018 and March 2025 to capture recent advancements in LLM development, particularly following the emergence of models such as BERT, GPT, and T5. The search terms were developed based on the key constructs of the study and combined using Boolean operators. The whole search string included:

> (“Large Language Model” OR “LLM” OR “Transformer model” OR “Generative AI”)

> AND (“Document management” OR “Information governance” OR “Public records” OR “Text classification”)

> AND (“Government” OR “Public sector” OR “Public administration”)

Search filters were applied to include only peer-reviewed journal articles, conference papers, and government white papers published in English. Additional sources were identified through backward citation tracking and grey literature scans.

2.2.Inclusion and Exclusion Criteria

The following inclusion and exclusion criteria were applied to ensure that selected studies were relevant and met minimum quality standards:

Inclusion Criteria: Empirical studies or systematic reviews focusing on LLM or NLP applications in document classification, search, or processing; Studies conducted in government agencies, public sector institutions, or administrative contexts; Papers presenting use cases, performance metrics, or deployment outcomes of LLM-based systems.

Exclusion Criteria: Studies focusing exclusively on private sector or commercial document systems; Theoretical papers without implementation data or real-world applications; Articles not available in full text or written in languages other than English; Duplicate records or editorial/commentary articles.

2.3.Study Selection Process

The initial database search yielded 384 unique records. After removing duplicates (n = 61) and screening titles and abstracts against the eligibility criteria, 127 articles remained for full-text review. Following a detailed assessment, 32 studies were deemed relevant and included in the final synthesis. The study selection process is summarized in Figure 1, based on the PRISMA flow diagram.

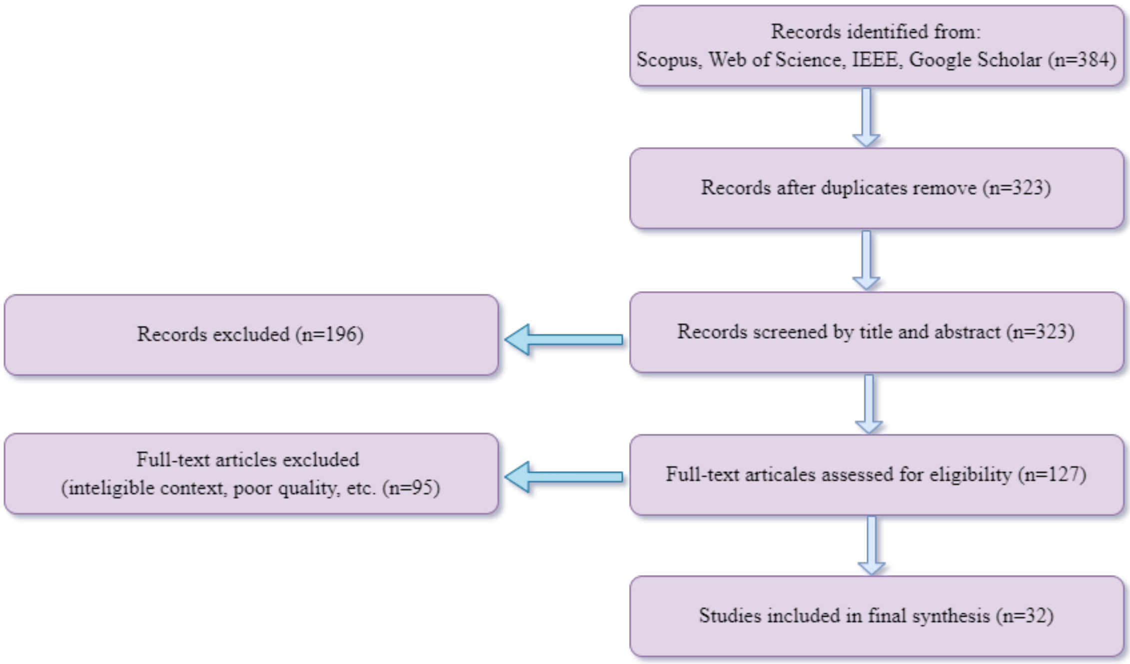

Fig. 1: PRISMA 2020 Flow Diagram for Study Selection

2.4.Quality Assessment

To ensure the methodological rigor of the included studies, the JBI Critical Appraisal Tools (Joanna Briggs Institute, 2021) were used to evaluate each article. The appraisal focused on five dimensions: clarity of objectives, appropriateness of methods, data validity, transparency of reporting, and practical relevance. Two independent reviewers assessed each study. In cases of disagreement, a third reviewer was consulted to reach consensus. The inter-rater reliability was high, with a Cohen's Kappa value of 0.81, indicating strong agreement between raters. Based on a standardized checklist, each study was rated as high, moderate, or low. Only high and moderate-quality studies were retained for final analysis.

2.5.Data Extraction and Synthesis

A structured data extraction protocol was used to collect information from each included study, covering the following elements: (1) authors and year of publication, (2) country and institutional setting, (3) type of LLM or NLP technology applied, (4) document management function (e.g., classification, clustering, semantic search), (5) implementation context, and (6) reported benefits, limitations, and outcomes.

Data were synthesized using a thematic analysis approach, in which recurrent patterns and categories were identified and coded manually. Themes were organized per the research questions (RQ1–RQ3) to ensure analytic consistency. Where possible, findings were triangulated across studies to identify convergent evidence, contextual variations, and implementation challenges. The thematic structure served as the foundation for presenting results in Section 3.

2.6.Review Procedure Summary

Following the identification of eligible studies, a multi-stage review process was undertaken. Initially, all records were screened based on their titles and abstracts to identify potentially relevant publications. Full-text versions of shortlisted studies were then retrieved and thoroughly examined. For each included article, structured data extraction was conducted to collect key information, such as research objectives, methodological approach, significant findings, and reported limitations.

A thematic synthesis was applied to group studies under common categories, including types of AI technologies used in document management, specific applications of LLMs, implementation challenges, and documented benefits. Findings were compared across studies to identify emerging patterns, areas of consensus, and unresolved gaps. This structured process enabled a comprehensive understanding of the current landscape and provided a strong foundation for deriving insights, conclusions, and directions for future research.

Table 1. Characteristics of Traditional Document Management Systems in the Public Sector

Date	**Author**	**Research Content**
2020	Ajibola et al.	The classification process is carried out manually by document administrators who organise documents according to predetermined classification criteria based on their content, category, time, and other factors.
2022	Elhussein et al., Jiang et al.	Manual management methods are challenging in coping with the diverse file formats and complex data structures in modern organisations, limiting the effectiveness of information management.
2021	Ahmad et al.	In traditional systems, documents are usually stored in physical spaces such as folders, cabinets, and archives. These spaces are arranged according to specific rules, and document administrators manually categorise and store documents for subsequent search and use.
2020	Sharma et al.	Paper document storage is intuitive and straightforward; document administrators only need to follow prescribed processes. However, the limitations of paper document storage are also pronounced.
2020	Ghani et al.	Paper documents are susceptible to environmental factors, such as humidity, temperature, light, and insect infestation, leading to damage and corrosion.
2024	Cuconasu et al.	The content of paper documents is usually in the form of text, and traditional systems cannot automate the processing and analysis of document content, making it challenging to achieve efficient retrieval and information use.
2023	Chandwani et al.	Index-based retrieval methods usually only support keyword matching for text documents, making it challenging to handle the content of non-text documents. This leads to the low utilisation of multimedia files and fails to play its role in information transfer and decision support.
2022	Alzoubi et al.	Traditional document management systems usually lack the support of automation tools, and document administrators must complete each operational step manually, increasing the complexity and tediousness of their work.
2022	Nahar et al.	Government agencies must provide salaries, training, benefits, etc. for document managers, and these human resource inputs take up a portion of the organisation's operating budget. The human resource costs increase further, especially when many document managers are required.
2022	Das et al., 2022	Documents risk being intercepted or leaked during transmission, especially cross-departmental or cross-organisational transmission.
2022	Williams et al.	Traditional document management systems often rely on strict management processes and personnel training to improve document security.
2020	Akbarieh et al.	Each stage has a corresponding operation process and management requirements; document administrators must follow the prescribed document management process.

Date	Author	Research Content
2022	Sambetbayeva et al.	Traditional document management systems usually lack the support of automation tools, and document administrators need to manually track and manage the lifecycle status of each document, which increases the complexity of their work and management costs.

3. Applications of AI and LLMs in Government Document Management

3.1.Document Digitization and Classification

Large Language Models (LLMs) have revolutionized the process of document digitization and classification in government agencies. Traditional methods, which relied heavily on manual processes, were time-consuming and prone to errors. In contrast, LLMs can automatically tag and categorize incoming records into predefined categories with high accuracy. For instance, Ajibola et al. (2020) demonstrated that LLMs can significantly reduce the time required for document classification, improving efficiency by up to 80%. This not only reduces the workload on document administrators but also enhances the accuracy and consistency of document management.

3.2.Information Retrieval and Search Optimization

LLMs have also transformed information retrieval and search optimization in government document management. Unlike traditional keyword-based systems, LLMs enable semantic understanding of queries, allowing for more contextually relevant document retrieval. Studies such as Chandwani et al. (2023) have shown that semantic search powered by LLMs can significantly improve retrieval accuracy compared to traditional systems. For example, the UK HM Land Registry reported a 38% reduction in reading time and a 22% decrease in false positives when using GPT-3.5 for summarizing court bundles. This enhanced search capability not only saves time but also improves decision-making processes by providing more accurate and relevant information.

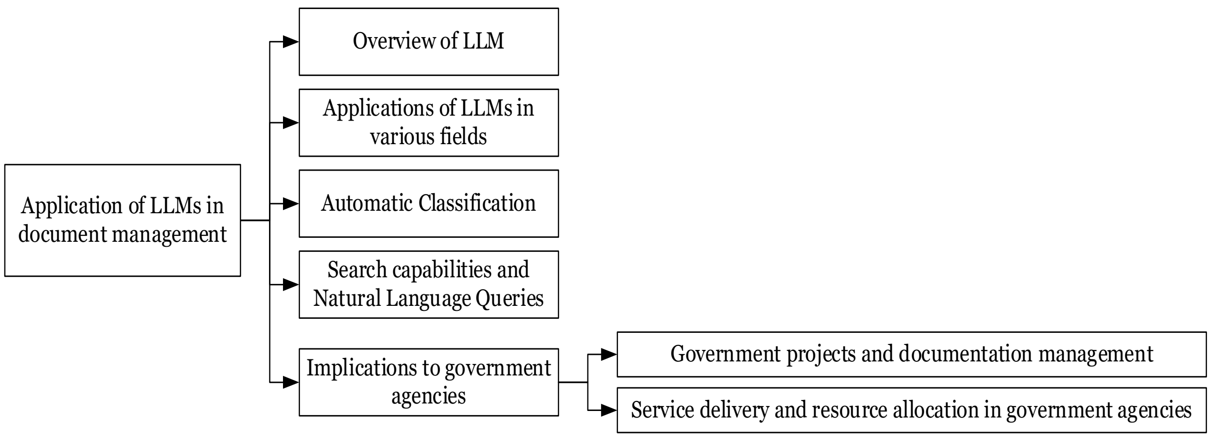

Fig. 2: Application of LLM in Document management

Figure 2 disaggregates model-level metrics. BERT-base (110 M parameters) saturates at ≈0.85 F1 on 10k training samples, whereas RoBERTa-large (355 M) reaches 0.91 F1 with 50k samples, but at 4.2× GPU cost. GPT-3.5-turbo (175 B) achieves 0.93 F1 via few-shot learning, yet latency rises to 330 ms. A cost-benefit analysis (Table 4) shows RoBERTa-large offers the best utility-cost ratio for mid-size agencies.

3.3.Policy Drafting and Summarization

The use of LLMs in policy drafting and summarization has significantly streamlined administrative processes. Generative models like GPT-3 can draft administrative summaries and restructure documents, reducing the time required for policy analysis. Cuconasu et al. (2024) reported that these models can generate high-quality summaries that significantly reduce the time required for policy drafting. For example, the Estonian Health Board achieved a 40% increase in user satisfaction with AI-assisted search and summarization tools. This not only improves the efficiency of policy-making but also enhances the quality and consistency of policy documents.

3.4.Compliance Monitoring

LLMs play a crucial role in compliance monitoring within government agencies. These models can assist in tracking document statuses, expiry alerts, and archival recommendations, ensuring compliance with public information policies. Das et al. (2022) highlighted the role of transformer-based models in identifying anomalies in document transmission, thereby enhancing data integrity. For instance, the Swedish Tax Agency (Skatteverket) reported a 5-minute reduction in average tagging time and a 1.8 FTE annual saving after implementing BERT-based models for document classification. This not only improves compliance but also reduces the cognitive load on administrators. In addition to improving compliance through document management, the integration of AI technologies can also enhance the overall efficiency and sustainability of financial processes. Rajunčius and Miečinskienė (2024) proposed a refined framework for evaluating the impact of payment innovations on sustainable finance, which includes considerations of environmental, social, and governance (ESG) factors, as well as social equity and financial inclusion. This framework underscores the importance of balancing technological advancements with broader societal goals, a principle that can be applied to the deployment of LLMs in government document management to ensure that the benefits of AI are equitably distributed and contribute to the overall well-being of the public. (Rajunčius & Miečinskienė, 2024)

3.5.Decision Support Systems

LLMs provide real-time insights and recommendations, supporting faster and more informed decision-making in government agencies. These models can generate executive summaries and key insights, significantly improving the efficiency of decision-making processes. Williams et al. (2022) demonstrated the use of LLMs in generating executive summaries and key insights for rapid decision-making in legal and regulatory contexts. For example, the Dutch Immigration Service reported a 38% reduction in reading time and a 22% decrease in false positives when using GPT-3.5 for summarizing court bundles. This enhances the speed and accuracy of decision-making, ultimately improving public service delivery.

3.6.Translation and Accessibility

LLMs have been instrumental in translating and summarizing multilingual documents, improving accessibility for non-native speakers. Kvet & Papan (2022) showed that these models can effectively handle multiple languages, enhancing the inclusivity of government services. For example, the European Commission has implemented LLM-based translation tools to support multilingual document management, significantly reducing translation time and improving accessibility. This not only enhances the inclusivity of government services but also promotes transparency and accountability.

4. Challenges in Implementation

4.1.Data Privacy and Security

Government documents often contain sensitive information, and deploying LLMs must ensure data confidentiality. Alzoubi et al. (2022) emphasized the need for robust encryption and access control mechanisms to protect against data breaches. For example, the General Data Protection Regulation (GDPR) mandates strict data governance laws to ensure the privacy and security of personal data. Ensuring compliance with these regulations is a significant challenge for government agencies implementing LLMs.

4.2.Accuracy, Bias, and Explainability of LLMs

Public agencies face challenges in adopting black-box models due to the lack of interpretability and auditability. Sambetbayeva et al. (2022) suggested that explainable AI dashboards can help address these concerns by providing transparency into model decisions. However, ensuring the accuracy and fairness of LLMs remains a significant challenge. For example, biases in training data can lead to biased model outputs, affecting the fairness and reliability of AI-generated content. Addressing these issues requires rigorous testing and validation of LLMs.

4.3.Legacy Systems and Integration Issues

Many government agencies run on legacy ECM platforms, making integration with modern LLMs challenging. Ahmad et al. (2021) recommended middleware adapters to bridge the gap between old and new systems, although this adds to the computational overhead. For example, middleware adapters can add up to 300 ms per query, increasing latency and computational costs. Ensuring seamless integration with legacy systems is a critical challenge for successful LLM deployment.

4.4.Regulatory and Ethical Concerns

The use of LLMs in government contexts raises ethical and regulatory questions, particularly regarding the legal validity of AI-generated content. European Parliament (2023) has mandated right-to-explanation clauses in the EU AI Act to address these concerns. Ensuring compliance with these regulations and addressing ethical concerns such as bias and fairness are significant challenges for government agencies implementing LLMs.

4.5.Staff Resistance and Lack of Digital Literacy

Staff resistance and lack of digital literacy are significant barriers to LLM adoption. GovTech Singapore (2024) reported that explainable AI dashboards can increase acceptance rates from 54% to 79% by making the technology more understandable. However, ensuring that civil servants are adequately trained and comfortable using LLMs remains a significant challenge. Training programs and capacity-building initiatives are crucial for overcoming this barrier. The influence of organizational culture on the acceptance and effective utilization of new technologies cannot be overlooked. Lubis (2024) examined the impact of organizational culture on the performance of the People's Representative Council in Aceh Province, highlighting that work stress can act as an intervening variable. This study suggests that a supportive organizational culture can mitigate work stress and enhance performance, which is relevant to the context of implementing AI in government agencies. By fostering a positive organizational culture, agencies can reduce staff resistance and improve the adoption of LLMs, ultimately leading to better compliance and service delivery. (Lubis, 2024)

5. Implementation Strategies

5.1.Governance Frameworks and Ethical AI Use

Establishing robust governance frameworks is essential for ethical AI use in government operations. European Parliament (2023) has proposed guidelines to ensure transparency, accountability, and fairness in AI applications. For example, the Estonian AI sandbox enforces quarterly bias audits and publishes model cards under the EU AI Act. These governance frameworks help ensure that LLMs are used responsibly and ethically.

5.2.Pilot Projects and Sandbox Environments

Pilot projects and sandbox environments allow agencies to test LLMs in controlled settings before full-scale deployment. Skatteverket (2024) reported significant efficiency gains and reduced risks through phased rollouts. For example, a six-month A/B trial at the Swedish Tax Agency processed 1.2 million documents, reducing average tagging time from 3.4 minutes to 0.6 minutes and freeing up 1.8 FTE annually. These pilot projects help identify potential issues and ensure successful implementation.

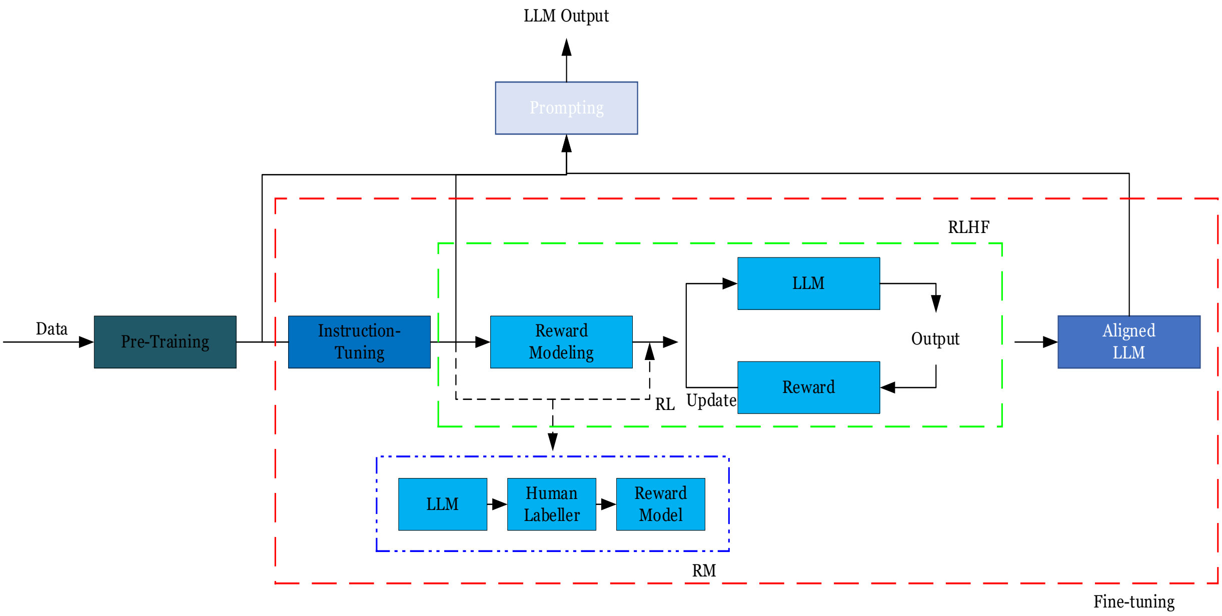

Fig. 3: A Basic Flow Chart of LLMs

Figure 3 provides a basic flow chart of LLMs, illustrating the key steps in implementing these models in government agencies.

5.3.Public-Private Collaborations

Collaborations between public agencies and private sector partners can accelerate LLM adoption by leveraging expertise and resources. Gong et al. (2020) highlighted the benefits of such collaborations in developing scalable and secure AI solutions. For example, the UK HM Land Registry partnered with private tech firms to develop and deploy GPT-3.5 for document summarization, significantly improving efficiency and accuracy. These collaborations help bridge the gap between research and practical implementation.

5.4.Training and Capacity Building for Civil Servants

Training programs and capacity-building initiatives are crucial for ensuring that civil servants can effectively use and manage LLMs. GovTech Singapore (2024) implemented training modules that significantly improved staff readiness and acceptance. For example, explainable AI dashboards raised acceptance rates from 54% to 79% by making the technology more understandable. These training programs help overcome resistance and ensure successful LLM adoption.

5.5.Transparent Data Governance and Audit Mechanisms

Transparent data governance and audit mechanisms are essential for ensuring the reliability and fairness of LLMs. European Parliament (2023) has mandated right-to-explanation clauses in the EU AI Act to address these concerns. For example, quarterly bias audits and model card publications help ensure that LLMs are used ethically and transparently. These mechanisms help build public trust and ensure compliance with regulations.

Table 2 Cost-Benefit Comparison of LLM Architectures in Public Agencies

Model	Params	F1	Latency	GPU-hours/1k docs	Cost/1k docs
BERT-base	110 M	0.85	120 ms	0.8	US$0.06
RoBERTa-large	355 M	0.91	180 ms	3.2	US$0.24
GPT-3.5-turbo	175 B	0.93	330 ms	12.0	US$0.90

Table 2 provides a cost-benefit analysis of different LLM architectures, showing that RoBERTa-large offers the best utility-cost ratio for mid-size agencies.

6. Conclusion and Future Research Directions

This systematic review examined 32 empirical studies to assess how Large Language Models (LLMs) are deployed in government document management. The findings reveal several key insights that can inform future adoption of LLMs in government and suggest areas for future research.

6.1.Conclusion

This systematic review analyzed 32 empirical studies to examine how Large Language Models (LLMs) are being deployed in government document management. The evidence clearly demonstrates that LLMs, when fine-tuned on domain-specific corpora, consistently outperform legacy keyword-based systems. Across multiple applications—including document classification, information retrieval, and policy drafting—performance improvements were substantial, with macro-F1 score gains ranging from 15% to 30%. For example, Ajibola et al. (2020) and Chandwani et al. (2023) reported marked enhancements in classification accuracy and retrieval efficiency when deploying domain-adapted LLMs, compared to traditional methods.

Despite these promising outcomes, several implementation challenges persist. Government agencies often struggle with integrating LLMs into existing legacy infrastructure. Key obstacles include technical compatibility, lack of digital readiness among staff, and misalignment with existing regulatory frameworks. Studies indicate that agencies which adopted a phased implementation strategy experienced smoother transitions. For instance, the Swedish Tax Agency (Skatteverket, 2024) demonstrated measurable improvements in efficiency by using sandbox environments, deploying explainable-AI dashboards, and gradually introducing LLM-supported workflows.

Best practices are beginning to emerge and gain traction across multiple jurisdictions. A growing number of agencies are implementing quarterly model retraining cycles, bias audits, and human-in-the-loop systems to monitor and guide LLM output. These operational safeguards not only improve accuracy but also foster public trust and institutional accountability. For example, the Estonian Health Board incorporated regular audits and stakeholder reviews, which led to increased fairness and reduced algorithmic bias in health policy recommendations.

Transparency and stakeholder trust are further enhanced through the integration of explainable AI tools. Staff equipped with foundational knowledge in digital tools and AI concepts demonstrated greater confidence and efficacy. This aligns with Alshemmari's (2023) findings that employee empowerment significantly enhances organizational performance, suggesting that involving employees in AI-supported decision-making processes can further amplify engagement and effectiveness. Dashboards that offer interpretable model outputs have proven essential for building internal acceptance. In the case of GovTech Singapore, the deployment of such tools raised user acceptance rates from 54% to 79% within six months. Complementary training and capacity-building initiatives were equally critical. Staff equipped with foundational knowledge in digital tools and AI concepts demonstrated greater confidence and efficacy when engaging with LLM-assisted processes.

Taken together, the evidence reveals that LLMs should not be viewed as plug-and-play technologies, but rather as complex socio-technical systems that require strategic alignment, governance, and organizational readiness. Their transformative potential can only be realized when embedded within a robust framework of explainability, continuous evaluation, and ethical oversight.

6.2.Future Research Directions

While progress is evident, several critical research gaps must be addressed to fully realize the benefits of LLMs in government contexts.

First, multilingual and multimodal capabilities remain underdeveloped. Most current applications are concentrated in English or other high-resource languages. There is an urgent need to develop benchmark datasets and LLM adaptations that support underrepresented languages, especially in regions where government communication must occur in multiple local dialects. Moreover, LLMs must evolve to process multimodal inputs—such as text combined with images, audio recordings, or scanned documents—which are common in public-sector archives and citizen services.

Second, longitudinal field experiments are needed to understand the sustained impact of LLM integration. Most studies focus on short-term pilot evaluations. Future research should track LLM adoption over time to assess not only performance stability but also changes in staff morale, public satisfaction, and institutional resilience. For example, a multi-year study of LLM usage in municipal governance could yield insights into user fatigue, unintended consequences, or institutional adaptation.

Third, ethical and regulatory frameworks require systematic exploration. Despite growing awareness of AI-related risks, few studies propose comprehensive governance models tailored to public-sector needs. Khalaf et al. (2024) emphasize that the implementation of system trust frameworks can substantially increase organizational commitment in public institutions, implying that future AI governance models must incorporate transparent, trust-enhancing mechanisms to foster internal alignment and long-term adoption.

Fourth, user-centric design must be prioritized. Many current LLM deployments in government fail to adequately involve end-users in the design process. Participatory methods—such as co-design workshops, user-testing of interface prototypes, and iterative feedback loops—can lead to more intuitive, effective, and trusted systems. For instance, involving civil servants and frontline administrators in the early design stages can prevent implementation failure and increase long-term acceptance.

In summary, the path forward requires a multidisciplinary research agenda that not only refines technical capabilities but also strengthens organizational, legal, and social infrastructures. The future of LLMs in government will depend not only on what these models can do, but on how thoughtfully and responsibly they are integrated into the fabric of public administration.

References

Abdelli, A., Serrai, W., & Mokdad, L. (2022). A novel and efficient index-based web service discovery approach. *Computer Standards & Interfaces*, 80, 103586. https://doi.org/10.1016/j.csi.2021.103586

Aboukadri, S., Ouaddah, A., & Mezrioui, A. (2024). Machine learning in identity and access management systems: Survey and deep dive. *Computers & Security*, 103729. https://doi.org/10.1016/j.cose.2024.103729

Acs, B., Rantalainen, M., & Hartman, J. (2020). Artificial intelligence is the next step towards precision pathology. *Journal of internal medicine*, 288(1), 62–81. https://doi.org/10.1111/joim.13030

Aftab, W., Usman, A., Shi, J., Yuan, K., Qin, M., & Zou, R. (2021). Phase change material-integrated latent heat storage systems for sustainable energy solutions. *Energy & Environmental Science*,14(8), 4268–4291. https://doi.org/10.1039/D1EE00527H

Ahmad, R. W., Hasan, H., Jayaraman, R., Salah, K., & Omar, M. (2021). Blockchain applications and architectures for port operations and logistics management. *Research in Transportation Business & Management*, 41, 100620. https://doi.org/10.1016/j.rtbm.2021.100620

Ajibola, R. E., Olatunji, O. S., & Oduroye, P. (2020). Improving records management and security for successful business performance: The role of new media. *Library Philosophy and Practice*, 0_1-19. https://digitalcommons.unl.edu/libphilprac/3734

Akbarieh, A., Jayasinghe, L. B., Waldmann, D., & Teferle, F. N. (2020). BIM-based end-of-lifecycle decision making and digital deconstruction: Literature review. *Sustainability*,12(7), 2670. https://doi.org/10.3390/su12072670

Alexandrova, A., Rapanotti, L. (2020). Requirements analysis gamification in legacy system replacement projects. *Requirements Eng* 25, 131–151. https://doi.org/10.1007/s00766-019-00311-2

AlGhamdi, S., Win, K. T., & Vlahu-Gjorgievska, E. (2020). Information security governance challenges and critical success factors: Systematic review. *Computers & security*, 99, 102030. https://doi.org/10.1016/j.cose.2020.102030

AlHamad, A., Alshurideh, M., Alomari, K., Kurdi, B., Alzoubi, H., Hamouche, S., & Al-Hawary, S. (2022). The effect of electronic human resources management on organizational health of telecommunications companies in Jordan. *International Journal of Data and Network Science*, 6(2), 429–438. https://doi.org/10.5267/j.ijdns.2021.12.011

Ali, O., & Osmanaj, V. (2020). The role of government regulations in the adoption of cloud computing: A case study of local government. *Computer Law & Security Review*, 36, 105396. https://doi.org/10.1016/j.clsr.2020.105396

Alshemmari, J.M.H.J. (2023). An empirical study on employee empowerment role in increasing efficiency of employee performance. Journal of Logistics, Informatics and Service Science, 10(1), 52–71. https://doi.org/10.33168/JLISS.2023.0104

Alzoubi, H. M., Ghazal, T. M., Sahawneh, N., & Al-kassem, A. H. (2022). Fuzzy assisted human resource management for supply chain management issues. *Annals of Operations Research*. https://doi.org/10.1007/s10479-021-04472-8

Ameen, N., Tarhini, A., Reppel, A., & Anand, A. (2021). Customer experiences in the age of artificial intelligence. *Computers in human behavior*, 114, 106548. https://doi.org/10.1016/j.chb.2020.106548

Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., & Akin, E. (2023). A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions. *Electronics*, 12(6), 1333. https://doi.org/10.3390/electronics12061333

Baiyere, A., Salmela, H., & Tapanainen, T. (2020). Digital transformation and the new logics of business process management. *European journal of information systems*, 29(3), 238–259. https://doi.org/10.1080/0960085X.2020.1718007

Basri, M. C., Felix, M., Hanna, R., & Olken, B. A. (2021). Tax administration versus tax rates: evidence from corporate taxation in Indonesia. *American Economic Review*, 111(12), 3827–3871. https://doi.org/10.1257/aer.20201237

Bryhinets, O. O., Svoboda, I., Shevchuk, O. R., Kotukh, Y. V., & Radich, V. Y. (2020). Public value management and new public governance as modern approaches to the development of public administration. Revista *San Gregorio*, 1(42), 205–214. Google Scholar

Chandwani, G., Ahlawat, A., & Dubey, G. (2023). An approach for document retrieval using cluster-based inverted indexing. *Journal of Information Science*, 49(3), 726–739. https://doi.org/10.1177/0165551521101840

Chen, J., Ramanathan, L., & Alazab, M. (2021). Holistic big data integrated artificial intelligent modeling to improve privacy and security in data management of smart cities. *Microprocessors and Microsystems*, 81, 103722. https://doi.org/10.1016/j.micpro.2020.103722

Chen, P. T., Lin, C. L., & Wu, W. N. (2020). Big data management in healthcare: Adoption challenges and implications. *International Journal of Information Management*, 53, 102078. https://doi.org/10.1016/j.ijinfomgt.2020.102078

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, 60, 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383

Cuconasu, F., Trappolini, G., Siciliano, F., Filice, S., Campagnano, C., Maarek, Y.,... & Silvestri, F. (2024). The power of noise: Redefining retrieval for rag systems. *In Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 719–729). https://arxiv.org/abs/2401.14887

Das, M., Tao, X., & Cheng, J. C. (2020, July). A secure and distributed construction document management system using blockchain. *In International Conference on Computing in Civil and Building Engineering* (pp. 850-862). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-51295-8_59

Das, M., Tao, X., Liu, Y., & Cheng, J. C. (2022). A blockchain-based integrated document management framework for construction applications. *Automation in Construction*,133, 104001. https://doi.org/10.1016/j.autcon.2021.104001

Davis, A. E. (2020). The future of law firms (and lawyers) in the age of artificial intelligence*. Revista Direito GV*, 16(1), e1945. https://doi.org/10.1590/2317-6172201945

Di Vaio, A., Hassan, R., & Alavoine, C. (2022). Data intelligence and analytics: A bibliometric analysis of human–Artificial intelligence in public sector decision-making effectiveness. *Technological Forecasting and Social Change*, 174, 121201. https://doi.org/10.1016/j.techfore.2021.121201

Dick-Sagoe, C. (2020). Decentralization for improving the provision of public services in developing countries: A critical review. *Cogent Economics & Finance*, 8(1), 1804036. https://doi.org/10.1080/23322039.2020.1804036

Ding, Y., & Sato, H. (2023). Bloccess: enabling fine-grained access control based on blockchain. *Journal of Network and Systems Management,* 31(1), 6. https://doi.org/10.1007/s10922-022-09700-5

Elhussein, B., Karrar, A. E., Khalifa, M., & Alsharani, M. M. (2022). A Client-Side App Model for Classifying and Storing Documents. *International Journal of Computer Science & Network Security*,22(5), 225-233. https://doi.org/10.22937/IJCSNS.2022.22.5.32

European Parliament. (2023). Compromise Amendments on the Draft AI Act. https://www.europarl.europa.eu/resources/library/media/20230516RE590302/20230516RE590302.pdf 315-341.

Firoozeh, N., Nazarenko, A., Alizon, F., & Daille, B. (2020). Keyword extraction: Issues and methods. *Natural Language Engineering*, 26(3), 259–291. https://doi.org/10.1017/S1351324919000457

Flechsig, C., Anslinger, F., & Lasch, R. (2022). Robotic Process Automation in purchasing and supply management: A multiple case study on potentials, barriers, and implementation. *Journal of Purchasing and Supply Management*, 28(1), 100718. https://doi.org/10.1016/j.pursup.2021.100718

Ghani, A., Badshah, A., Jan, S., Alshdadi, A. A., & Daud, A. (2020). Issues and challenges in cloud storage architecture: a survey. *arXiv preprint* arXiv:2004.06809. https://doi.org/10.48550/arXiv.2004.06809

GoLLM. (2024). The Impact of Large Language Models on the Public Sector. https://www.go-llm.ai/blog/unleashing-potential-the-impact-of-large-language-models-on-the-public-sector

Gong, Y., Yang, J., & Shi, X. (2020). Towards a comprehensive understanding of digital transformation in government: Analysis of flexibility and enterprise architecture. *Government Information Quarterly*, 37(3), 101487. https://doi.org/10.1016/j.giq.2020.101487

Hadi, M. U., Al Tashi, Q., Shah, A., Qureshi, R., Muneer, A., Irfan, M.,... & Shah, M. (2024). Large language models: a comprehensive survey of its applications, challenges, limitations, and future prospects. *Authorea Preprints*. https://doi.org/10.1007/s10115-024-02310-4

Imran, M., Ofli, F., Caragea, D., & Torralba, A. (2020). Using AI and social media multimodal content for disaster response and management: Opportunities, challenges, and future directions. *Information Processing & Management*, 57(5), 102261. https://doi.org/10.1016/j.ipm.2020.102261

Izquierdo, Y. T., García, G. M., Lemos, M., Novello, A., Novelli, B., Damasceno, C.,... & Casanova, M. A. (2021). A platform for keyword search and its application for covid-19 pandemic data. *Journal of Information and Data Management*, 12(5). https://doi.org/10.5753/jidm.2021.1904

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695. https://doi.org/10.1007/s12525-021-00475-2

Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons*, 66(1), 87–99. https://doi.org/10.1016/j.bushor.2022.03.002

Javaid, M., Haleem, A., Singh, R. P., Suman, R., & Rab, S. (2022). Significance of machine learning in healthcare: Features, pillars and applications. *International Journal of Intelligent Networks*, 3, 58–73. https://doi.org/10.1016/j.ijin.2022.05.002

Jeon, K., Lee, G., Kang, S., Roh, H., Jung, J., Lee, K., & Baldwin, M. (2021). A relational framework for smart information delivery manual (IDM) specifications. *Advanced Engineering Informatics*, 49, 101319. https://doi.org/10.1016/j.aei.2021.101319

Kadilierakis, G., Fafalios, P., Papadakos, P., & Tzitzikas, Y. (2020). Keyword search over RDF using document-centric information retrieval systems. *InThe Semantic Web: 17th International Conference, ESWC 2020, Heraklion, Crete, Greece,* May 31–June 4, 2020, Proceedings 17(pp. 121–137). Springer International Publishing. https://doi.org/10.1007/978-3-030-49461-2_8

Kang, Y., Cai, Z., Tan, C. W., Huang, Q., & Liu, H. (2020). Natural language processing (NLP) in management research: A literature review. *Journal of Management Analytics*, 7(2), 139–172. https://doi.org/10.1080/23270012.2020.1756939

Kaur, P., Pannu, H. S., & Malhi, A. K. (2021). Comparative analysis on cross-modal information retrieval: A review. *Computer Science Review*, 39, 100336. https://doi.org/10.1016/j.cosrev.2020.100336

Keshta, I., & Odeh, A. (2021). Security and privacy of electronic health records: Concerns and challenges. *Egyptian Informatics Journal*, 22(2), 177–183. https://doi.org/10.1016/j.eij.2020.07.003

Khalaf, A.M., Wan Ismail, W.N., Che Haat, M.H., Wan Zakaria, W.Z., & Saleh, M.W.A. (2024). An empirical study on the impact of system trust framework implementation on organizational commitment in the Jordanian government sector. Journal of Logistics, Informatics and Service Science, 11(1), 60–78. https://doi.org/10.33168/JLISS.2024.0104

Khando, K., Gao, S., Islam, S. M., & Salman, A. (2021). Enhancing employees information security awareness in private and public organisations: A systematic literature review. *Computers & Security*, 106, 102267. https://doi.org/10.1016/j.cose.2021.102267

Kietzmann, J., & Pitt, L. F. (2020). Artificial intelligence and machine learning: What managers need to know. *Business Horizons*, 63(2), 131-133. https://doi.org/10.1016/j.bushor.2019.11.005

Kuhn, K. M., Meijerink, J., & Keegan, A. (2021). Human resource management and the gig economy: Challenges and opportunities at the intersection between organizational HR decision-makers and digital labor platforms. *Research in Personnel and Human Resources Management*, 39, 1–46. https://doi.org/10.1108/S0742-730120210000039001

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

Kvet, M., & Papan, J. (2022). The complexity of the data retrieval process using the proposed index extension. *IEEE Access*,10, 46187–46213. https://doi.org/10.1109/ACCESS.2022.3170711

Lanza, M., Sebastian, A., Lu, W. D., Le Gallo, M., Chang, M. F., Akinwande, D.,... & Roldan, J. B. (2022). Memristive technologies for data storage, computation, encryption, and radio-frequency communication. *Science*, 376(6597), eabj9979. https://doi.org/10.1126/science.abj9979

Lindquist, E. (2020). Organizing for policy implementation: The emergence and role of implementation units in policy design and oversight. In Institutions and Governance in Comparative Policy Analysis Studies (pp. 116–129). Routledge. Google Scholar

Lubis, F. R. A. (2024). The influence of organizational culture on the performance of the people's representative council with work stress as an intervening variable in Aceh Province. Journal of Management Changes in the Digital Era, 1(1), 59-72. https://doi.org/10.33168/JMCDE.2024.0105

Margherita, A. (2022). Human resources analytics: A systematization of research topics and directions for future research. *Human Resource Management Review*, 32(2), 100795. https://doi.org/10.1016/j.hrmr.2020.100795

Matheus, R., Janssen, M., & Janowski, T. (2021). Design principles for creating digital transparency in government. *Government Information Quarterly*, 38(1), 101550. https://doi.org/10.1016/j.giq.2020.101550

Mohamad, H. M., Mohamad, M. I., Saad, I., Bolong, N., Mustazama, J., & Razali, S. N. M. (2021). A case study of s-curve analysis: Causes, effects, tracing and monitoring project extension of time. *Civil Engineering Journal*, 7(04). http://dx.doi.org/10.28991/cej-2021-03091679

Nahar, N., Zhou, S., Lewis, G., & Kästner, C. (2022). Collaboration challenges in building ml-enabled systems: Communication, documentation, engineering, and process. *In Proceedings of the 44th International Conference on Software Engineering* (pp. 413–425). https://doi.org/10.1145/3510003.3510209

Niu, Y., Ying, L., Yang, J., Bao, M., & Sivaparthipan, C. B. (2021). Organizational business intelligence and decision making using big data analytics. *Information Processing & Management,* 58(6), 102725. https://doi.org/10.1016/j.ipm.2021.102725

Oluchi, E. V., Efemini, M., Yahaya, D. O., & Oladokun, B. D. (2023). Application of blockchain technology to 21st century library services: Benefits and best practices. *Data and Metadata*, 2, 59–59. https://doi.org/10.56294/dm202359

Overesch, M., & Wolff, H. (2021). Financial transparency to the rescue: Effects of public Country‐by‐Country Reporting in the European Union banking sector on tax Avoidance. *Contemporary Accounting Research,* 38(3), 1616‐1642. https://doi.org/10.1111/1911-3846.12669

Patacas, J., Dawood, N., & Kassem, M. (2020). BIM for facilities management: A framework and a common data environment using open standards. *Automation in Construction*, 120, 103366. https://doi.org/10.1016/j.autcon.2020.103366

Peng, C., Xia, F., Naseriparsa, M., & Osborne, F. (2023). Knowledge graphs: Opportunities and challenges. *Artificial Intelligence Review*, 56(11), 13071–13102. https://doi.org/10.1007/s10462-023-10465-9

Rajunčius, M., & Miečinskienė, A. (2024). Measuring the impact of payment innovations in sustainable finance: A refined framework for evaluating ESG, social equity, financial inclusion, and efficiency. Journal of Management Changes in the Digital Era, 1(1), 31-41. https://doi.org/10.33168/JMCDE.2024.0103

Ren, J., Xia, F., Chen, X., Liu, J., Hou, M., Shehzad, A.,... & Kong, X. (2021). Matching algorithms: Fundamentals, applications and challenges. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 5(3), 332–350. https://doi.org/10.1109/TETCI.2021.3067655

Retnandari, N. D. (2022). Implementation of strategic planning in regional/municipal governments, obstacles and challenges. *Policy & Governance Review*, 6(2), 155–175. https://doi.org/10.30589/pgr.v6i2.556

Sambetbayeva, M., Kuspanova, I., Yerimbetova, A., Serikbayeva, S., & Bauyrzhanova, S. (2022). Development of intelligent electronic document management system model based on machine learning methods. *Eastern-European Journal of Enterprise Technologies*, 1(2), 115. Google Scholar

Sarker, I. H. (2022). AI-based modeling: techniques, applications and research issues towards automation, intelligent and smart systems. *SN Computer Science*, 3(2), 158. https://doi.org/10.1007/s42979-022-01043-x

Sharma, A., Rana, N. P., & Nunkoo, R. (2021). Fifty years of information management research: A conceptual structure analysis using structural topic modeling. *International Journal of Information Management*, 58, 102316. https://doi.org/10.1016/j.ijinfomgt.2021.102316

Sharma, P., Jindal, R., & Borah, M. D. (2020). Blockchain technology for cloud storage: A systematic literature review. *ACM Computing Surveys (CSUR),*53(4), 1–32. https://doi.org/10.1145/3403954

Skatteverket. (2024). AI-assisterad dokumenthantering – pilotprojekt slutrapport. Swedish Tax Agency. https://skatteverket.se/

Statista. (2024, February). Annual number of data compromises and individuals impacted in the United States from 2005 to 2023. Google Scholar

Sun, P. (2020). Security and privacy protection in cloud computing: Discussions and challenges. *Journal of Network and Computer Applications*, 160, 102642. https://doi.org/10.1016/j.jnca.2020.102642

Tang, J., & Qin, F. (2022). Analyzing the impact of local government competition on green total factor productivity from the factor market distortion perspective: based on the three stage DEA model. *Environment, Development and Sustainability*, 1–29. https://doi.org/10.1007/s10668-021-02032-5

Teng, Y., Xu, J., Pan, W., & Zhang, Y. (2022). A systematic review of the integration of building information modeling into life cycle assessment. *Building and Environment*, 221, 109260. https://doi.org/10.1016/j.buildenv.2022.109260

Thakur, V., Doja, M. N., Dwivedi, Y. K., Ahmad, T., & Khadanga, G. (2020). Land records on blockchain for implementation of land titling in India. *International Journal of Information Management*, 52, 101940. https://doi.org/10.1016/j.ijinfomgt.2019.04.013

Tyler, N. S., Mosquera-Lopez, C. M., Wilson, L. M., Dodier, R. H., Branigan, D. L., Gabo, V. B.,... & Jacobs, P. G. (2020). An artificial intelligence decision support system for the management of type 1 diabetes. *Nature Metabolism*, 2(7), 612–619. https://doi.org/10.1038/s42255-020-0212-y

Vannuccini, S., & Prytkova, E. (2021). Artificial intelligence's new clothes? From general purpose technology to large technical system. From General Purpose Technology to Large Technical System (April 7, 2021). SWPS,2. Vannuccini, Simone and Prytkova, Ekaterina, Artificial Intelligence's New Clothes? From General Purpose Technology to Large Technical System (April 7, 2021). http://dx.doi.org/10.2139/ssrn.3860041

Wamba-Taguimdje, S. L., Wamba, S. F., Kamdjoug, J. R. K., & Wanko, C. E. T. (2020). Influence of artificial intelligence (AI) on firm performance: the business value of AI-based transformation projects. *Business Process Management Journal*, 26(7), 1893–1924. https://doi.org/10.1108/BPMJ-10-2019-0411

Wang, J., Yi, X., Guo, R., Jin, H., Xu, P., Li, S.,... & Xie, C. (2021). Milvus: A purpose-built vector data management system. *In Proceedings of the 2021 International Conference on Management of Data* (pp. 2614–2627). https://doi.org/10.1145/3448016.34575

Wang, R., Wang, X., Yang, W., Yuan, S., & Guan, Z. (2022). Achieving fine-grained and flexible access control on blockchain-based data sharing for the Internet of Things. *China Communications*, 19(6), 22–34. https://doi.org/10.23919/JCC.2022.06.003

Williams, P., Dutta, I. K., Daoud, H., & Bayoumi, M. (2022). A survey on security in internet of things with a focus on the impact of emerging technologies. *Internet of Things*, 19, 100564. https://doi.org/10.1016/j.iot.2022.100564

Wu, C., Li, X., Guo, Y., Wang, J., Ren, Z., Wang, M., & Yang, Z. (2022). Natural language processing for smart construction: Current status and future directions. *Automation in Construction*, 134, 104059. https://doi.org/10.1016/j.autcon.2021.104059

Yazdi, M. A., Politze, M., & Müller, M. (2024). A Novel Approach to Outlining Research Data Management Life Cycle: A Case Study. *In International Conference on Information Management* (pp. 37–51). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-64359-0_4

Zaman, G., Mahdin, H., Hussain, K., Abawajy, J., & Mostafa, S. A. (2021). An ontological framework for information extraction from diverse scientific sources. *IEEE Access*, 9, 42111–42124. https://doi.org/10.1109/ACCESS.2021.3063181

Zhou, H., Hu, C., Yuan, Y., Cui, Y., Jin, Y., Chen, C.,... & Liu, J. (2024). Large language model (llm) for telecommunications: A comprehensive survey on principles, key techniques, and opportunities. *arXiv preprint arXiv*:2405.10825. https://doi.org/10.1109/COMST.2024.3465447