© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716115

A Software Architecture Model for Reducing Citizen–Government Trust Deficit in AI-Driven e-Government

LAWAL, K. H[1], OJERINDE O. A[2], ALENOGHENA I. B[3], FOLUSO AYENI[4]
[1, 2, 3] *Software Engineering Department Federal University of Technology Minna, Nigeria*
[4] *Department of Management Information Systems, Metro State University, USA*

Abstract- ***The integration of AI into many platforms allows for automation and robotic Process Automation (RPA), predictive analytics, and data-driven decision-making, each of which streamlines operations. Despite the progress, citizen trust in AI-empowered government systems is fragile due to issues of algorithmic opacity, data privacy vulnerabilities, bias and weak accountability mechanisms. While existing research addresses trustworthy AI broadly through ethical principles, governance frameworks, and regulatory approaches, little has been written on operationalizing trust in software architecture. This gap is currently addressed by the this paper, which proposes a software architectural model intended to mitigate the citizen–government trust deficit in AI-enabled e-government systems. Through a Design Science Research approach, trust-related requirements are derived from literature and transformed into architectural entities integrated within distinct layers of the system. This paper proposes a novel framework that combines transparency-by-design, explainable AI, accountability services and data governance with citizen engagement approaches. The paper makes an important contribution by showing that trust can be baked into the system as a fundamental architectural feature rather than considered entirely at the policy or governance level.***

Keywords: Artificial Intelligence (AI), E-Government Systems, Citizen Trust, Software Architecture, Explainable AI (XAI), Transparency-by-Design, Accountability Mechanisms, Data Governance and Privacy, Algorithmic Bias and Fairness, Citizen Engagement and Contestability.

I. INTRODUCTION

Digital transformation is one of the defining features of modern public administration. Governments worldwide are progressively utilizing e-government systems to automate public service delivery, control administrative processes, and interact with citizens in more effective and scalable manners (United Nations, 2022; Janssen, Charalabidis & Zuiderwijk, 2021). Over the years, these systems have transformed into multifaceted online platforms that go beyond merely providing information; they now encompass a wide range of functionalities essential for transactional purposes, inter-agency collaboration and real-time decision-making. AI in recent years has proved a central enabling technology of this transformation, through smart automation, predictive analytics, natural language processing and algorithmic decision support (Sun & Medaglia 2019; Wirtz et al. 2023).

E-government systems utilising AI appear fit to translate these advantages into meaningful net gains for societies, in the form of speedier service delivery, lower operational costs, more accurately targeted policies and increased administrative efficiency. Emerging fields that reflect the use of AI in government are migrant automated welfare eligibility assessment, intelligent tax administration, virtual assistants for citizen services, and predictive resource allocation (Wirtz, Weyerer & Geyer 2019; Eom & Lee (2022). But with such benefits have come existing challenges around trust in government institutions exacerbated by the arrival of AI. For many citizens, automated decision-making systems appear opaque, intrusive and hard to contest when such a system determines their access to public services, legal rights or social benefits (Fonfría et al. 2023).

Trust is one of the main pillars of good governance. For the e-government trust is a precursor for citizen acceptance of numerous aspects as willingness adopts digital services, gives personal data and accepts decisions that are made or assisted by automated systems Abdul et al., (2020). In the absence of trust, even technically advanced systems may see low usage, resistance or public backlash. Apart from augmenting existing trust concerns such as ‘black box’ type issues relating to machine learning models, algorithmic bias and discrimination, uncertainty over what is being

done with our data and a lack of accountability when an algorithm goes wrong or causes harm (Burrell, 2016; Ribeiro et al., 2020; Winfield & Jirotka, 2022), AI technologies also add in complexities–dimensionalities interwoven within the governance of society.

Most of the current debate on trust-worthy AI in government centres on ethical principles, legal requirements and top-level governance arrangements (Floridi et al., 2018; Morselli & Zuiderwijk, 2026). Although these approaches are relevant, they do not in themselves provide a complete solution. If the systems themselves are opaque and unaccountable by design, then trust will not be gained through policy decrees or post hoc oversight mechanisms. We draw attention and lay out the case for embedding trust directly by design into the technical architecture of AI-enhanced e-government systems. Trust is first-class architectural concern — by treating it as such, governments can operationalize (or graph and measure) transparency, accountability, and citizen participation within the structural organization of digital public services. Hence the aim of this study is to formulate a software architecture model that systematically mitigates the citizen–government trust deficit in AI-driven e-government systems.

II. BACKGROUND AND RELATED WORK

Empirical research on e-government adoption has consistently shown that trust is among the most important predictors of citizen acceptance and continued use of digital public services (building on work by: Cole et al. 2011; Jansen et al. 2019; Kim, Chiu & Teyar Headi, Seong-Won, Hee-Chan 2020). Most early studies recognized trust drivers including: system quality; service efficiency; and institutional credibility. Further studies went beyond focus on the technologies themselves to highlight how perceived risk, privacy protection or transparency plays a role in the establishment of citizen trust (Carter & Bélanger, 2005; Al-Saqaf & Nielsen, 2019; Rahmadany & Mansyur, 2021). These studies collectively highlight that technology capability on its own is not enough — citizens need to view governments' systems as trustworthy and fair, in alignment with democracy and societal values.

The advent of artificial intelligence in the public sector has created a new layer of complexity to this discourse on trust. Unlike the traditional information systems, where data is transformed based on a well-defined algorithm, AI-driven systems are often built on top of complex models that implement non-linear operations whose internal logic cannot be easily understood even by developers. The inability to explain these decisions has sparked fears about procedural fairness and due process, especially in domains where automated decisions can have significant social, economic or legal implications (Burrell, 2016; Abdul et al., 2020; Ribeiro et al., 2020). Scholars have also noted risks from biased training data and algorithmic design decisions, leading to unfair outcomes and eroding public trust in automated government decision-making Aninze & Bhogal, (2024).

To address these issues, an increasing number of trustworthy and ethical AI literature has proposed principles such as transparency, explicability, fairness, accountability, human oversight (Floridi et al., 2018; Jobin et al., 2020; Winfield & Jirotka, 2022). In particular, these values manifest in the mechanisms of governance through policy instruments, ethical frameworks and regulatory projects such as the Organisation for Economic Co-operation and Development's (OECD) AI Principles and the European Union's AI Act. Nevertheless, these contributions are at best normative and lack traction to produce implementation.

A complementary perspective provided by software architecture research is its focus on the way high level system design decisions impact critical quality attributes, including security, reliability, maintainability, scalability to do accountability (Bass, Clements & Kazman 2013; Garlan 2021). Architectural decisions define how system components interact, control data flows, and exercise controls and oversight. However, few studies explicitly address trust as an architectural property in AI-driven e-government systems. Most existing architectures favor interoperability and speed, as well as trust through external governance arrangements or human supervisory mechanisms.

This gap indicates that architecture-oriented solution for trustworthy AI in e-government is missing. Embedding trust-enhancing mechanisms at the system architecture level will allow us to work beyond mere symbolic transparency toward operational transparent, accountable and citizen-empowering systems. By integrating findings from the literature on e-government adoption, scholarship on trustworthy AI and principles of software architecture theory, this study sets forth a singular architectural model that embeds trust directly into the fabric of AI-powered government systems.

III. PROBLEM ANALYSIS AND SYSTEM REQUIREMENTS

3.1 Problem Analysis: Sources of Trust Deficit in AI-Driven E-Government

Trust deficit between citizens and AI-based e-governance systems is a problem due to technical, institutional, and socio-cognitive reasons. While for traditional information systems, data is mainly used to build profiles of individuals and communities based on which decisions are made; AI-based systems often perform inferential and predictive tasks that will have direct effects on 'the public', determining the eligibility of citizens towards certain projects or assessing risks and prioritizing of services. Moreover, when these type of decisions are seen as being non-transparent or immune to challenge this can cause citizens to see them as arbitrary or unjust, leading to a breakdown in trust (Burrell, 2016; Gong, Kordzadeh and Nelson 2023).

Thus, one of the key sources of trust deficit is algorithmic opacity. Many AI models—especially those employing deep learning techniques—are opaque so citizens cannot know how inputs are transformed into outcomes. In public-sector settings, such opacity contradicts democratic values including transparency, due process, and the right to explanation (Ribeiro et al., 2020; Janssen et al., 2021).

A second challenge concerns data governance and privacy. AI-driven e-government systems often leverage cross-agency, big sData types of aggregation. Citizens frequently lack insight into the processes by which their data are acquired, handled, transmitted, or stored and can be subjected to surveillance (Bajracharya 2024). Poor data governance policies further exacerbate these issues and erode institutional trust.

The third source of trust deficit is algorithmic bias and perceived injustice. Studies have shown that biases in training data or the design of ML models can lead to pervasive disparities between social groups Aninze & Bhogal, (2024). Such outcomes are especially damaging in government contexts, because they can reinforce social inequalities and lower perceptions of legitimacy and fairness.

Finally, the lack of accountability is an enduring problem. When A.I. systems inform or automate public decisions, who is responsible for mistakes or harm the developer of the system, the agency deploying it or the algorithm? (Busuioc, 2021). As this diffusion of responsibility undermines institutional accountability, it further diminishes citizens' trust in redress mechanisms [5].

3.2 Trust-Oriented System Requirements

To address these challenges, trust must be translated into explicit system requirements rather than treated as an abstract governance ideal. Based on the problem analysis and existing trustworthy AI literature, this study identifies the following core trust-oriented requirements for AI-driven e-government systems:

1) Transparency Requirement
   The system must provide visibility into decision processes, data flows, and system behavior. Transparency should be operationalized through system logs, traceable decision pipelines, and user-facing explanations rather than policy statements alone (Garlan, 2021).
2) Explainability Requirement
   AI-driven decisions affecting citizens must be explainable in forms appropriate to different stakeholders. This includes technical explanations for auditors and simplified, human-understandable explanations for citizens Alshehri & Drew, (2021).
3) Fairness and Bias Mitigation Requirement
   The system must incorporate mechanisms for detecting, monitoring, and mitigating bias throughout the AI lifecycle, including data collection, model training, and deployment (Aninze & Bhogal, (2024).

4) Accountability Requirement
Clear responsibility attribution must be embedded within system design. Decision logs, audit trails, and human-in-the-loop controls are required to support oversight and redress (Ottun & Flores, 2025).
5) Data Governance and Privacy Requirement
The architecture must enforce data minimization, purpose limitation, access control, and compliance with data protection regulations to enhance institutional trust (Hjerppe et al., 2019).
6) Citizen Participation and Contestability Requirement
Citizens must have mechanisms to question, contest, and appeal AI-driven decisions. Participation should be supported through feedback channels and procedural safeguards embedded in the system (Al-Saqaf & Nielsen, 2019).

These requirements form the foundation for the proposed software architecture model developed in later sections.

IV. RESEARCH METHODOLOGY

4.1 Research Approach: Design Science Research

A Design Science Research (DSR) approach is utilized in this study, as it is suitable for creating artifacts that can solve complex socio-technical issues. The goal of DSR is the intentional design and assessment of artifacts which may include models, frameworks or architectures that aim at solving a specific organizational or societal problem (Hevner et al., 2004).

Citizen–government trust deficit for the use of AI in e-government is a socio-technical problem involving complex interactions among technology, institutions, and humans. Standard empirical approaches lack the closure to deal with this complexity, since they usually aim at explaining phenomena rather than designing interventions. Since DSR presents a structured approach for transforming theoretical knowledge into practical system designs, it is fit for purpose in this study.

4.2 Design Science Process

The research followed a structured DSR process consisting of the following stages:

1) Problem Identification and Motivation
The trust deficit in AI-driven e-government systems was identified through an extensive review of literature on e-government adoption, trustworthy AI, and public-sector digital governance. The motivation for the study lies in the lack of architecture-level solutions addressing trust.
2) Objective Definition for a Solution
Based on the identified problems, the study defined objectives for a trust-centric solution, including transparency-by-design, explainability, accountability, and citizen engagement.
3) Design and Development
A software architecture model was designed to embed trust-enhancing mechanisms across multiple system layers. The architecture integrates technical and institutional controls to operationalize trust requirements.
4) Demonstration
The proposed architecture is demonstrated conceptually by mapping trust requirements to architectural components and explaining how these components address identified trust challenges.
5) Evaluation
The architecture is evaluated analytically using established trust, transparency, and governance criteria derived from recent literature (Gong et al., 2023; Winfield & Jirotka, 2022).
6) Communication
The results are communicated through this paper to both academic and practitioner audiences,

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716115

contributing to research on AI-driven e-government and trustworthy system design.

4.3 Methodological Contribution

This study makes a methodological contribution by showing how high level trust principles can be instantiated into low level architectural artifacts, through the application of Design Science Research to resolve trust in AI-driven e-government. While entirely conceptual discussions of ethical AI remain quite valuable, the DSR approach also intuitively narrows in on what we can design for: trust as a system property, which is the core claim made by this paper.

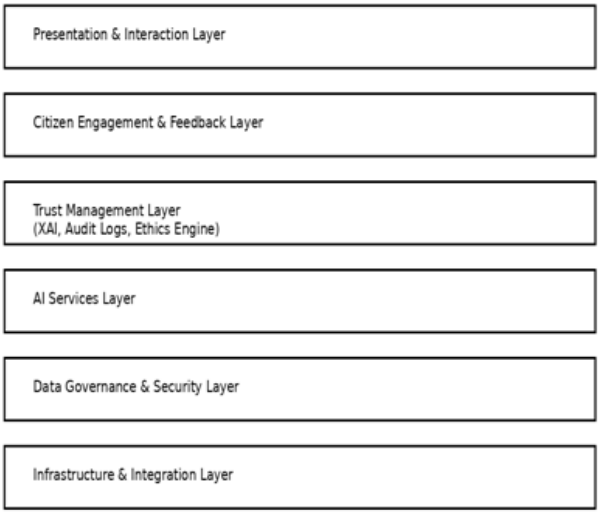
V. PROPOSED SOFTWARE ARCHITECTURE MODEL

5.1 Architectural Design Rationale

The central premise of this study is that building citizen trust in AI-driven e-government systems cannot come strictly from governance frameworks, but has to be incorporated directly into the design of those systems. The appropriate level of abstraction for operationalizing trust is found in software architecture, which is the study of how system components cooperate to produce decisions and what mechanisms exist to control and monitor them (Bass et al., 2013; Garlan, 2021).

This section introduces the proposed architecture which aims to fulfill the trust-oriented system needs presented in Section 3, through the integration of technical, organizational and citizen-facing mechanisms under a single framework. Instead of treating trust as an external property assessed post-deployment, the architecture embeds mechanisms for its elevation throughout various layers—making transparency, accountability, and contestability emergent properties of the system.

Figure 1: Trust-Centric AI-Driven E-Government Architecture



5.2 Architectural Overview

The proposed model adopts a layered software architecture consisting of six interrelated layers:

1) Presentation and Interaction Layer
2) Citizen Engagement and Contestability Layer
3) Trust Management Layer
4) AI Services Layer
5) Data Governance and Security Layer
6) Infrastructure and Integration Layer

Thus, while each layer addresses certain trust issues, together they contribute to ensuring overall trustworthy behavior. Trust is defined as a cross-cutting architectural quality feature, which is especially focused in the Trust Management Layer where mediates interactions between AI-components, data resources and citizen-facing services.

5.3 Presentation and Interaction Layer

Citizens and public officials interact with e-government services through interfaces provided by the Presentation and Interaction Layer. These interfaces can take the form of web portals, mobile applications or conversational agents. This layer is responsible not just for usability and accessibility, but also how transparent system decisions are communicated from a trust perspective.

Results produced by AI delivery to citizens come with context, confidence scores, and explanations of how the decisions can be challenged or appealed. Research

by Abdul et al (2020) and Ribeiro et al (2020) has shown that even when user does not agree with what the system has decided, explainable content enough to trust the reasons..

5.4 Citizen Engagement and Contestability Layer
This layer makes democratic principles operative by facilitating citizen participation, feedback, and contestability. It includes provisions for submitting queries, complaints, requests for review and appeals against AI-generated decisions.

Unlike traditional feedback modules, which operate outside the core logic of a system, this layer is architecturally embedded into decision pipelines. This allows citizen input to initiate review processes, human oversight, or system audits. Filling accountability gaps and fortifying institutional trust (Al-Saqaf & Nielsen, 2019; Winfield & Jirotka, 2022) offers the potential for embedding contestation at an architectural level.

5.5 Trust Management Layer (Core Layer)
The central contribution of the proposed architecture is the Trust Management Layer. It is a layer between the AI services, data resources and end-user applications. Its key functions include:

- Explainability services
- Bias and fairness monitoring
- Decision logging and audit trails
- Ethics and compliance enforcement

(Vimosh and Zhang, 2023) Explainability services are designed to provide different types of explanations (post hoc vs. intrinsic) to stakeholders. Continuous bias monitoring modules check outputs for discrimination patterns, facilitating early detection and intervention [4].

In traceability, decision logs and audit trails facilitate external audits and internal reviews for accountability. This resonates with the emerging recommendations for trustworthy AI governance in the public sector (Gong et al., 2023).

5.6 AI Services Layer
The AI Services Layer is the place where machine learning models and analytical engines for prediction, classification, and decision support are run. We choose models to deploy in this layer with trust considerations as a priority, trading off interpretability, robustness and auditability for performance-driving metrics.

Where high complexity models are required, complementary explainability techniques are employed to lessen opacity. This design choice illustrates a broader consensus that performance optimization in isolation is no longer enough for public sector AI systems (Sun & Medaglia, 2019).

5.7 Data Governance and Security Layer
This level ensures security, permissioning, and legal compliance. It ensures quality of the data, tracking its provenance and secure sharing of the data across agencies. The importance of strong data governance is critical to maintain the trust of society, especially in AIbased environments that are highly reliant on large-scale data integration (Janssen et al., 2021)

VI. AI AND TRUST-ENABLING MECHANISMS

6.1 Explainable Artificial Intelligence
Explainable AI (XAI) mechanisms provide stakeholders insight on how and why AI systems arrive at certain outcomes. XAI should be used as a service rather than just an optional addition, ensuring only one interface to generate explanations across applications Alshehri & Drew (2021).

6.2 Fairness, Bias Mitigation, and Monitoring
Such mechanisms have a role in each aspect of the AI lifecycle, from data preprocessing through deployment to post-deployment monitoring. Ongoing fairness evaluation increases legitimacy, and diminishes probabilities of systemic inequity Aninze & Bhogal, (2024).

6.3 Accountability and Human Oversight
You have human-in-the-loop controls to augment, not replace, human judgment in high-risk decisions. Bajracharya (2024) cites clear escalation paths and attribution of responsibility as critical to building institutional accountability.

VII. EVALUATION FRAMEWORK AND CRITERIA

7.1 Purpose of Evaluation
Unlike the role of evaluation in purely empirical studies, it has a different purpose in design-oriented research. Instead of empirically verifying hypotheses, the aim is to evaluate whether or not the suggested artifact successfully solves the identified problem and meets the specified requirements (Hevner et al., 2004). This study evaluates if the proposed software architecture is a solution to eliminating sources of citizen–government trust deficit in AI-powered e-government systems.

Due to the conceptual nature of architecture, its evaluation is analytical through a criteria-based framework built from the existing literature on trust, transparency, and trustworthy AI in the public sector (Gong et al., 2023; Winfield & Jirotka, 2022).

7.2 Evaluation Dimensions
The proposed architecture is evaluated across five primary dimensions:
7.2.1 Transparency and Explainability
This dimension measures how well the architecture enables visibility into AI-led decision loops. Some key metrics for the same might be availability of explainable AI services, service traceability on decision pipelines, user-surfacing expository visualizations, and so on. Architectures that offer grounded explanations and auditability of decisions are more likely to improve process trust and sense of fairness (Alshehri & Drew,2021).

The proposed model meets this requirement by embedding explainability services into the Trust Management Layer, where explanations can be continuously provided and consulted through applications.

7.2.2 Accountability and Auditability
Accountability assessment looks at what can be traced, inspected and challenged about the behavior of a system. This includes action logs, audit trails and clear responsibility attribution mechanisms. In public sector contexts accountability is a requirement for legitimacy and redress Bajracharya, (2024).

Because accountability is key, embed audit logging and compliance monitoring into the architecture as an integral part of it instead of optional add-ons.

7.2.3 Fairness and Bias Management
It assesses whether the architecture allows for detection and mitigation of bias along the AI lifecycle. Ongoing fairness checks and bias assessments are essential to maintain trust in automated public choices Aninze &Bhogal, (2024.

This need will be satisfied by the proposed architecture through specialized bias monitoring modules connected to AI services and decisions outputs.

7.2.4 Data Governance and Privacy Protection
Evaluation in this dimension weighs whether data collection, processing and sharing are governed by explicit process controls. Robust data governance lowers perceptions of surveillance and misuse, which can have a positive impact on institutional trust (Janssen et al., 2021).

A dedicated layer takes care of access control, data provenance, and legal compliance to enforce data governance across the architecture.

7.2.5 Citizen Participation and Contestability
This dimension measures whether citizens are given meaningful avenues to engage with, and question or challenge, AI-driven decisions. Democratic e-participation and contestability underpins democratic accountability and trust (Al-Saqaf & Nielsen, 2019).

This architecture meets this requirement as it embeds citizens' engagement and appeal mechanisms in the decision workflow.

7.3 Summary of Evaluation

Overall, the assessment shows that the architecture proposed meets the fundamental trust-based requirements determined in previous sections. By embedding these principles within the design of each system, the architecture addresses key drivers of citizen–government trust deficit in AI-driven e-government environments.

VIII. DISCUSSION

8.1 Theoretical Implications

This research enhances trust models in the architectural domain as an addition to e-government and information systems literature. Although previous literature has treated trust mostly as a psychological or institutional concept, this paper provides evidence that trust can also be understood as a designable system property.

This research is unique in that it attempts to integrate findings from normative discussions of trustworthy AI, and either literature or frameworks around software architecture, presenting a middle ground between theorization surrounding ethical AI and the building of systems. Thus, this relates to the ongoing discussions on how democratically oriented values can become manifest and be implemented in digital government infrastructures.

8.2 Practical Implications

For policymakers, it constitutes a principled architecture with which accountability needs can be designed into public AI systems from day one, instead of having to rely on oversight after the fact. This methodology enables proactive governance and minimizes the likelihood of public backlash.

To system architects and developers, the model provides reusable architectural patterns for adding suitable mechanisms to implement explainability, auditability and citizen engagement aspects. We can adapt these patterns across many domains in government without fundamentally redesigning them.

8.3 Limitations

This study has several limitations. First, the proposed architecture is theoretical and has not been empirically validated in real-world deployment. Second, the assessment is based on analytical reasoning instead of experimental testing. Taken together, the architecture does not represent all contextual constant variations in political, legal and cultural environments.

Due to these limitations, future research is needed toward empirical validation and contextual adaptation.

IX. CONCLUSION AND FUTURE RESEARCH

In this paper, we introduced a software architecture model for mitigating the citizen–government trust deficit in AI-driven e-government systems. The study made it possible to operationalize concepts such as transparency, explainability, accountability, fairness and citizen participation within system design by treating trust as a first-class architectural concern.

The main contribution of this research is the policy shift from high level principles on trustworthy AI in government to architectural solutions. A text-based approach to understanding the legitimation of AI systems in e-government contexts could contribute to further design approaches for AI-enhanced governmental tools.

Future work should focus on validating the proposed architecture through empirical pilot implementations and case studies in diverse government domains. Further studies could also analyze quantitative aspects of trust increase and be interested in how architectural trust mechanism interplayed with the organizational and regulatory factors.

REFERNCES

[1] Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2020). Trends and trajectories for explainable, accountable and intelligible systems. *PACM HCI*, 4(CSCW3).

[2] Al-Saqaf, W., & Nielsen, R. (2019). Trust and transparency in e-government systems. *Government Information Quarterly*, 36(4), 101386.

[3] Alshehri, M., & Drew, S. (2021). *Government Information Quarterly*

[4] Aninze, A., & Bhogal, J. (2024). Artificial Intelligence Life Cycle: The Detection and Mitigation of. In Proceedings of the International Conference on AI Research. Academic Conferences and publishing limited.

[5] Bajracharya, K. (2024). Big Data and Artificial Intelligence Integration in Modernizing Governance and Public Administration Practices. Global Research Perspectives on Cybersecurity Governance, Policy, and Management, 8(12), 34-47.

[6] Bass, L., Clements, P., & Kazman, R. (2013). Software Architecture in Practice (3rd ed.). Addison-Wesley.

[7] Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. Big data & society, 3(1), 2053951715622512.

[8] Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. Public administration review, 81(5), 825-836.

[9] Carter, L., & Bélanger, F. (2005). The utilization of e-government services: citizen trust, innovation and acceptance factors. Information systems journal, 15(1), 5-25.

[10] de Cisneros Fonfría, J. J. J., Amador, A. M. G., Gonzalez, A. Q., & Fernandez, L. P. (2023). Investigating the T-stub connection with different web-to-flange joint configurations. Engineering Structures, 294, 116715.

[11] Eom, S. J., & Lee, J. (2022). Digital government transformation in turbulent times: Responses, challenges, and future direction. Government Information Quarterly, 39(2), 101690.

[12] Garlan, D. (2021). Architectural tactics for trustworthy AI systems. *Journal of Systems and Software*, 173, 110864.

[13] Gong, B., Kordzadeh, N., & Nelson, K. (2023). Trustworthy AI in public services: A systematic review and research agenda. *Government Information Quarterly*, 40(2), 101683.

[14] Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75–105.

[15] Hjerppe, K., Ruohonen, J., & Leppänen, V. (2019, September). The general data protection regulation: requirements, architectures, and constraints. In 2019 IEEE 27th International Requirements Engineering Conference (RE) (pp. 265-275). IEEE.

[16] Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2021). Data governance in digital government: Revisiting the role of trust. *Government Information Quarterly*, 38(3), 101585.

[17] Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature machine intelligence, 1(9), 389-399.

[18] Morselli, F., & Zuiderwijk, A. (2026). Open science in academia: a framework for monitoring universities' Open Science programs. Research Evaluation, 35, rvaf058.

[19] Ottun, A. R. O., & Flores, H. (2025). Trustworthy AI in Practice: A Comprehensive Review of Human Oversight and Human-in-the-Loop Approaches. Authorea Preprints.

[20] Rahmadany, A. F., & Mansyur, A. (2021). The Implementation E-Government to Increase Democratic Participation: The Use of Mobile Government. Jurnal Studi Sosial dan Politik (JSSP), 5(1), 22-34.

[21] Ribeiro, M. T., Singh, S., & Guestrin, C. (2020). "Why should I trust you?" Explaining the predictions of any classifier. *ACM SIGKDD*.

[22] Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. Government information quarterly, 36(2), 368-383.

[23] Winfield, A. F., & Jirotka, M. (2022). Ethical governance is essential to building trust in AI systems. *Philosophical Transactions of the Royal Society A*, 380(2214).

[24] Winfield, A. F., van Maris, A., Salvini, P., & Jirotka, M. (2022). An ethical black box for social robots: a draft open standard. arXiv preprint arXiv:2205.06564.

[25] Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector—applications and challenges. International Journal of Public Administration, 42(7), 596-615.

[26] Wirtz, J., Kunz, W. H., Hartley, N., & Tarbit, J. (2023). Corporate digital responsibility in service firms and their ecosystems. Journal of Service Research, 26(2), 173-190.