Journal of Artificial Intelligence and AI Ethics

Research Article

Open Access Full Text Article

From Hierarchies to Loops: Rethinking Public Sector Structures for Responsible AI Integration

Sebastain Obeta[1*], Adoration Chigere[2], Dr Ikpe Ibanga[3], Ruth Itua[4], Linda Oraegbunam[5], Andrew Chinonso Nwanakwaugwu[6], Godson Nmesoma Ozioma[7], Chidinma Anumaka[8]

[1]*Cambridge University UK.*
[2]*University of Oulu, Finland.*
[3]*School of Health, Leeds Beckett University.*
[4]*Security Operations Analyst; Epaton LTD.*
[5]*Data Science Researcher Sycom Labs.*
[6]*Accenture UK.*
[7]*University of Salford.*
[8]*University of Nottingham.*

***Correspondence:**

Sebastain Obeta
Cambridge University UK.

Received: November 21, 2025; **Published:** January 09, 2026

How to cite this article:

Obeta, S., Chigere, A., Ibanga, I., Itua, R., Oraegbunam, L., Nwanakwaugwu, A. C., Ozioma, G. N. & Anumaka, C. (2026), 'From hierarchies to loops: Rethinking public sector structures for responsible AI integration', *Journal of Artificial Intelligence and AI Ethics*, vol. 1, no. 1, pp. 1–10.

Abstract

As artificial intelligence (AI) systems shift from automating tasks to supporting cognitive decision-making, public-sector organisations face both structural and philosophical disruptions. Traditional departmental hierarchies, designed to manage human limitations in complexity, are increasingly misaligned with AI-native workflows, which prioritise speed, feedback, and dynamic abstraction. This paper introduces the concept of "Neural Government Design," a framework where public organisations function as adaptive cognitive systems with looped information flows, recursive decision structures, and hybrid human-AI reasoning. Drawing on organisational theory, systems thinking, and real-world examples from the NHS and local councils, we examine how AI integration challenges traditional bureaucracy. We then present both hypothetical and existing use cases where looped, feedback-driven models yield more responsive, ethical, and efficient public services. Finally, we propose design principles and metrics for assessing intelligence throughput, collaborative yield, and cognitive compression in AI-enabled government structures. This paper aims to bridge academic theory and public sector transformation, offering actionable insights for researchers, technologists, and civil servants shaping the future of AI governance.

1. Introduction

The institutions we have were not designed for the intelligence we now possess.

Artificial intelligence (AI) is fundamentally transforming the public sector, marking a significant shift from the automation of routine tasks to complex, abstract decision-making processes (Mellouli *et al.*, 2024). Over the past decade, governments worldwide have increasingly invested in AI, reflecting both its technological maturation and its perceived potential to enhance efficiency, responsiveness, and public value. Over the past decade, governments worldwide have increasingly invested in AI, reflecting both its technological maturation and its perceived potential to enhance efficiency, responsiveness, and public value (Maalla, 2021; O'Connor *et al.*, 2024; Prakash *et al.*, 2023; Zhang *et al.*, 2022). This transformation, however, challenges the suitability of long-standing bureaucratic structures to fully realise AI's benefits while avoiding its risks.

Historically, governance has relied on "street-level bureaucrats," individuals exercising discretion in policy implementation under conditions of complexity and uncertainty (Bullock, 2019). The advent of information and communication technologies (ICT) enabled a shift to "screen-level" bureaucracies, where computer systems and databases routinised many decisions (Bovens and Zouridis, 2002). Today, the rise of AI is ushering in "systems-level" bureaucracies, in which ICT tools increasingly augment or even replace expert judgment in more sophisticated, high-stakes contexts (Bullock, 2019).

Early governmental adoption emphasised automation as a means to "do more with less" (Maciejewski, 2017), promising to streamline operations, reduce costs, and free human resources for higher-value tasks (Maalla, 2021). Yet, the reality has been more complex. A consistent gap persists between optimistic expectations and actual outcomes (Henriksen and Blond, 2023; O'Connor *et al*., 2024). While AI offers efficiency gains, its integration has raised serious concerns over accountability, transparency, bias, privacy, and job displacement (Alberti, 2019; Mellouli et al., 2024; Prakash et al., 2023). High-profile failures such as Australia's "Robodebt" scandal (Alberti, 2019; Russell and Norvig, 2022) (Alberti, 2019; Peter, 2023) and the Netherlands' System Risk Indication (SyRI) (O'Connor *et al*., 2024) demonstrate the risks of deploying AI without a deep understanding of its limitations, the specific problem it seeks to address, or the intended policy outcome. When automation is driven primarily by "managerial interests in cost-efficiency" and "political agendas of rationalisation and modernisation" (Henriksen and Blond, 2023), it can erode human capabilities and generate negative societal impacts (Alberti, 2019; Hartmann and Wenzelburger, 2021).

At the same time, AI capabilities are evolving beyond task automation toward generative reasoning, contextual abstraction, and strategic foresight functions once considered exclusive to human cognition (Kolt *et al*., 2025; Lubana *et al*., 2024). Such systems often behave like complex adaptive systems, displaying non-linear growth, cascading effects, and self-reinforcing feedback loops (Cen *et al*., 2024; Fu *et al*., 2022), where changes in one part can trigger unpredictable consequences across the network (Gao *et al*., 2022). Yet, public sector AI deployments frequently prioritize the needs of executives and managers focused on cost rationalization over the requirements of frontline practitioners and service users (Henriksen and Olesen, 2021), producing what (Henriksen and Blond, 2023) call a "parasitic symbiosis" that undermines the human-centred or intelligence augmentation (IA) ideal of amplifying human agency through human–machine collaboration (Sundar, 2020).

Legacy organisational structures are increasingly inadequate in this context and designed for linear workflows, departmental silos, and human-only cognition; they cannot scale intelligence in the ways advanced AI systems require. As AI shifts from automating discrete tasks to supporting high-level cognitive functions, it exposes a widening misalignment between technological capability and institutional form. Addressing this misalignment is not merely an operational challenge; it is a theoretical gap in organisational science. Existing models rarely account for the feedback-rich, adaptive, and non-linear properties of AI-enabled decision systems. Without new design principles, institutions risk embedding powerful AI into outdated structures, limiting its potential, amplifying unintended consequences, and constraining its capacity to enhance collective intelligence.

This paper responds to that gap by introducing the term and conceptual framework Neural Business Design (NBU), which reimagines organisations as cognition-centred systems built on feedback loops, cross-functional flows, and hybrid reasoning models. In doing so, it extends organisational theory to accommodate AI-native operational logic and proposes metrics for evaluating and scaling intelligence within operations. This contribution not only offers a new vocabulary for describing AI-integrated institutions, but also provides a foundation for replacing departmental silos in the emerging post-automation economy.

2. Background and Literature Context

The trajectory of organisational design in government has historically revolved around a central tension: the drive for efficiency through structured bureaucracy versus the growing imperative for adaptability, inclusivity, and human-centred governance. From Max Weber's foundational principles to contemporary responses to uncertainty and complexity, this evolution reflects both theoretical development and empirical recalibration in the face of changing governance realities.

2.1 The Evolution of Organisational Design in Government

Max Weber's conception of the "ideal type" of rational bureaucracy remains a cornerstone in organisational theory. It was not intended as a literal prescription but rather as an analytical framework to understand administrative order (Sager and Rosser, 2021; Yılmaz and Telsaç, 2021). His model emphasised formal hierarchy, clearly defined roles, rule-based operations, and meritocratic professionalism designed to achieve maximal efficiency and control in service delivery (Doğan, 2020; Hughes, 2014). However, Weber foresaw the dangers of this system becoming overly rigid and dehumanising. His notion of the "iron cage" captures the potential ossification of bureaucratic structures, their resistance to innovation, and detachment from the lived realities of citizens (Bouckaert, 2023).

Critics have long noted the empirical inadequacies of Weber's model, particularly its neglect of informal dynamics within organisations, the influence of external environments, and the complex motivations of human actors. Early sociologists challenged Weber's underestimation of political influence, informal norms, and the relational elements of bureaucratic life (Sager and Rosser, 2021). These critiques paved the way for more adaptive and context-sensitive theories of organisational design.

Jay Galbraith's Organisational Information Processing Theory (OIPT) extended organisational thought into the realm of contingency, emphasising the need for structures that match task complexity and uncertainty. Galbraith's model proposed strategies such as slack resources, self-contained units, enhanced information systems, and lateral relations to deal with growing informational demands (Dwivedi *et al*., 2021; Haußmann *et al*., 2012). While OIPT made valuable contributions to understanding organisational adaptation, it has been critiqued for its limited engagement with ambiguity and its tendency to assume that technology reduces uncertainty when in practice, it may exacerbate it (Cooper and Wolfe, 2005).

Henry Mintzberg further diversified organisational theory by proposing a taxonomy of structural configurations, including machine bureaucracy, professional bureaucracy, adhocracy, and others, based on contextual factors such as size, strategy, and technical environment (Ljevo and Šunje, 2021). Mintzberg's framework emphasised the interplay between coordination mechanisms and organisational complexity without prescribing a one-size-fits-all model.

Building on Mintzberg, recent innovations such as the Decentralised Science Pyramid Framework (DSPF) for Decentralised Autonomous Organisations (DAOs) have emerged, reconceptualising hierarchy in favour of technologically enabled, community-driven governance. The DSPF integrates Mintzberg's emphasis on coordination with decentralised, transparent mechanisms and aligns with contemporary organisational needs for flexibility, collaboration, and regulatory compliance (Weidener *et al*., 2024). However, DAOs also introduce novel challenges ranging from slow consensus-building and strategic fragmentation to accountability and security vulnerabilities (Wright, 2020).

This shift in organisational design becomes especially salient when public agencies operate in what Snowden categorises as "complex" or "chaotic" domains characterised by high uncertainty, interdependence, and non-linearity (Doğan, 2020). The inadequacies of traditional bureaucratic models in such environments are magnified with the integration of Artificial Intelligence (AI), which introduces further layers of complexity and unpredictability (Hughes, 2014).

Contemporary AI systems mirror complex adaptive systems in their emergent, non-linear, and cascading behaviours (Kolt *et al*., 2025). While AI promises enhanced efficiency and decision-making speed, its application in public governance has revealed substantial implementation gaps. The "sociology of expectations" illustrates how AI's promise often diverges from practice, obscuring social and technical complexities (O'Connor *et al*., 2024). Furthermore, empirical studies suggest that AI is frequently introduced not as a tool for human enhancement but for managerial efficiency, deepening concerns about dehumanisation, exclusion, and executive-centred design (Henriksen and Blond, 2023; Rességuier and Rodrigues, 2020).

This techno-managerial orientation can lead to a paradigm shift from a "professional treatment model" that emphasises contextual fairness and procedural justice to a bureaucratic rationality model driven by data-matching and automation (Hughes, 2014). In such settings, algorithms often perceived as objective embed normative choices that shape social outcomes, frequently without adequate transparency or accountability. Emerging human-centred paradigms argue for a more nuanced approach. (Ford, 2022) proposes a humanity-driven public administration that elevates emotion, perception, and lived experience. This approach reconfigures traditional value hierarchies, privileging public acceptance, social equity, and local contextuality over abstract metrics of effectiveness. In complex public domains, legitimacy and effectiveness are inseparable from public trust and perceived fairness. As (Overman and Schillemans, 2022) argue, accountability must be redefined in relational terms, especially for professional actors who do not conform to rigid hierarchical models.

While Weber's ideal type remains foundational, it is no longer sufficient. Galbraith and Mintzberg introduced models for managing complexity, while the advent of AI has made clear that public administration must evolve toward systems that are simultaneously adaptive, technologically informed, and deeply human-centric. Governance in the 21st century must balance efficiency with ethical and contextual responsiveness, designing institutions that are not only functionally effective but also socially legitimate.

2.2 AI in Public Services

The integration of AI into public services has been widely heralded for its transformative potential, yet the reality of its implementation reveals persistent tensions with entrenched legacy systems, administrative inertia, and systemic under-resourcing. While AI offers the promise of automation, prediction, and enhanced service delivery, its full potential remains constrained by historical and infrastructural path dependencies.

One of the most cited impediments is the incompatibility of AI with outdated IT infrastructure across many public institutions. For example, within the NHS and other government bodies, legacy systems often lack the flexibility to integrate (Adelekan *et al*., 2024; González García *et al*., 2019; Li, 2022). In the NHS Southwest London case, AI deployment was delayed by concurrent Picture Archiving and Communication System (PACE) upgrades and resource constraints in IT departments, ultimately leading to a selection process driven by ease of integration rather than performance (Shelmerdine *et al*., 2024).

Paradoxically, AI tools may intensify existing bottlenecks. Enhanced diagnostic accuracy can increase demand for downstream interventions, as seen in the rising volume of CT scans triggered by AI triage tools, placing additional strain on already stretched radiology departments. Thus, technological optimisation in one area may displace pressures to another, undermining net system efficiency (Shelmerdine *et al*., 2024).

Furthermore, bureaucratic and regulatory hurdles, particularly concerning data protection and research ethics, slow the integration of AI systems. Complex data governance frameworks and bespoke service level agreements frequently delay project timelines and reduce innovation flexibility (Adelekan *et al*., 2024). The practical difficulties of coordinating AI across organisational silos and departments compound these challenges.

Beyond operational issues, ethical concerns about AI's role in decision-making are paramount. In socially sensitive areas such as welfare or healthcare, removing human oversight raises questions about fairness, accountability, and public trust (Alshahrani *et al*., 2024). Transparency in algorithmic decision-making remains elusive, especially in black-box models, and the risks of discriminatory outcomes or the erosion of procedural justice remain high.

These challenges call for a holistic reform approach: upgrading technological infrastructure, cultivating AI literacy and workforce skills, designing transparent governance frameworks, and addressing organisational resistance. Without these complementary strategies, AI's transformative potential in the public sector will remain largely theoretical.

The integration of sophisticated AI technologies into public sector operations is consistently hindered by the prevalence and incompatibility of legacy IT systems (Adelekan *et al*., 2024; Shelmerdine *et al*., 2024). Public sector organisations, including the NHS and tax administrations, often possess outdated and inconsistent IT infrastructures that are not designed for seamless integration with modern AI tools (González García *et al*., 2019; Li, 2022). Similarly, in the UK healthcare, the strategic rollout of AI in Southwest London faced delays due to concurrent PACS (Picture Archiving and Communication System) upgrades across the network, which diverted IT staff attention and postponed integration decisions (Shelmerdine *et al*., 2024). The eventual selection of an AI product for an NHS trial was ultimately dictated by its ease of integration into existing IT and PACS frameworks, highlighting the practical constraints imposed by legacy systems. Furthermore, many NHS IT departments are severely under-resourced, limiting their capacity to undertake unfunded service evaluations and necessary updates that would facilitate AI deployment (Shelmerdine *et al*., 2024).

This dynamic poses a significant risk to systems such as the NHS, which already operate under restrictive funding and limited resources (González García *et al*., 2019). Preliminary findings from an AI trial in Southwest London, for instance, revealed a rise in patients receiving expedited CT examinations. While this acceleration benefits patient outcomes, it simultaneously increases the workload for radiology departments, potentially straining their capacity (Shelmerdine *et al*., 2024).

Moreover, the implementation of AI tools within the NHS is frequently hindered by bureaucratic hurdles and delays. The process of introducing new technologies or initiating research studies often encounters extensive and complex data protection requirements, which can significantly slow down project setup (Shelmerdine *et al*., 2024). These administrative challenges not only impede operational efficiency but also complicate the management of unique service level agreements (SLAs) and the integration of AI-derived insights into core business functions (Adelekan *et al*., 2024).

Finally, ethical concerns surrounding human exclusion from decision-making processes remain paramount. In domains such as complex social services, the feasibility and implications of removing human oversight raise critical questions (Alshahrani *et al*., 2024). Ensuring human involvement is essential for maintaining public trust, transparency, and fairness in AI-driven decisions. The deployment of AI in public services is indeed marked by a complex interplay of conformance to legacy structures, the emergence of decision-making bottlenecks, critical interpretability gaps, and systemic coordination issues. Addressing these interwoven challenges requires a holistic strategy encompassing technological upgrades, comprehensive training and skill development, robust governance frameworks emphasising transparency and accountability, and a proactive approach to managing organisational and cultural resistance.

Without such an integrated efforts, the transformative potential of AI to enhance public service delivery and decision-making will remain significantly constrained.

2.3 Sociotechnical and Cognitive Systems

The sociotechnical tradition in organisational theory offers vital insights into the integration of AI and emerging technologies into public systems. Trist and Emery's foundational research demonstrated that technological efficiency cannot be achieved in isolation from the social systems that mediate its use. Their concept of joint optimisation, where neither technical nor social systems are fully optimised in isolation, remains deeply relevant (Stanton, 2022). Similarly, Herbert Simon's theory of bounded rationality fundamentally redefined the understanding of human decision-making. Recognising the cognitive limitations of administrators, Simon emphasized "satisficing" rather than optimising and introduced the concept of procedural rationality to structure decision-making (Schwarz *et al*., 2022). His pioneering work in AI envisioned computers as cognitive aids, expanding the decision-making capacity of humans within complex environments (Thorstad, 2024).

Karl Weick's sense-making theory further enriches this perspective by highlighting the interpretive and emotional dimensions of organisational life. Organisations are not merely procedural machines but environments of narrative construction, emotional contagion, and retrospective meaning-making, particularly salient in periods of ambiguity or crisis (Cristofaro, 2022; Kundra and Dwivedi, 2023).

Together, these perspectives suggest a new vision for public organisations, not as static procedural engines but as adaptive cognitive platforms. AI, when integrated thoughtfully, has the potential to augment human cognition, extend bounded rationality, and facilitate new forms of sense making. However, this requires more than technological insertion. It demands systemic transformation.

This transformation must address affective responses to AI, not just rational evaluation. Emotional schemata shape acceptance, resistance, and organisational culture. Leaders, as sense givers, play a pivotal role in guiding collective interpretation and trust-building (Cristofaro, 2022). Trust, as (Thornton *et al*., 2022) argue, is a multidimensional construct shaped by reputation, shared values, and technological reliability. Public institutions must become "alchemists of trust," adapting trust models to specific contexts and systems.

Transitioning from hierarchies to loops involves dismantling rigid, top-down structures in favour of reflexive, iterative, and context-aware governance frameworks. These adaptive platforms must engage deeply with human cognition, emotion, and meaning-making while maintaining institutional safeguards of transparency and equity.

3. Theoretical Framework: Neural Government Design

The traditional organisational paradigm, defined by departmental divisions based on functional specialisation, has long constrained comprehensive insight and efficient decision-making. This fragmentation of information, often referred to as "departmental logic" stems from the historical evolution of technology, the specialisation of departmental roles, and legacy acquisition patterns (Steele, 2025). As a result, data becomes trapped in incompatible formats, disparate semantic structures, and incongruent security models. The operational consequences are significant: duplicate data entry, inconsistent reporting, and delays in strategic decision-making, particularly when holistic analysis requires manual integration across siloed systems.

3.1 Collapse of Departmental Logic

Public organisations are especially vulnerable to these inefficiencies due to their formal, rigid structures that prioritise control and exploitation over the flexibility needed for innovation (Selten and Klievink, 2024). Fragmented expertise across departments further impedes knowledge diffusion and retention, making adaptation to rapid technological change even more challenging.

Artificial intelligence (AI) is fundamentally reshaping this landscape by transcending the cognitive and structural limitations inherent in human-centric, functionally siloed approaches to data management and analysis (Bhima *et al*., 2023; Yildirim *et al*., 2022). AI systems are uniquely positioned as transformative agents, capable of ingesting and synthesising vast quantities of disparate data across traditionally separated functions. This capability enables organisations to overcome persistent challenges in capturing, transferring, and utilising institutional knowledge effectively across silos (Sira, 2024).

The mechanisms through which AI achieves this cross-functional integration are multifaceted. AI technologies act as a connective infrastructure, a kind of digital connective tissue linking previously isolated data environments and knowledge repositories. Unlike traditional integration methods that rely on rigid schema mapping, machine learning algorithms serve as adaptive bridges, detecting patterns, establishing relationships, and generating unified views across heterogeneous data sources. Through APIs and micro services, AI can also serve as architectural bridges in cloud environments, facilitating seamless data exchange and semantic harmonisation.

Beyond integration, AI accelerates routine tasks through automation and generates predictive insights from historical data, enabling organisations to anticipate trends and take proactive measures (Bhima *et al*., 2023). This not only liberates human resources for strategic thinking but also transforms data fragmentation from a liability into a source of innovation and competitive differentiation.

In essence, AI enables a shift from a fragmented, function-bound operational logic to an integrated, insight-driven paradigm. For example, an AI system trained on both health and housing data can identify systemic poverty patterns more rapidly and accurately than separate departments coordinating manually. This collapse of departmental logic reflects a deeper cognitive transformation: departments exist because human cognition is limited, and we divide work by function to manage scope and specialisation. AI, however, can ingest and synthesise across these boundaries, offering a more holistic and dynamic understanding of complex societal issues.

3.2 From Hierarchies to Loops

Traditional public sector structures, often characterised by a vertical flow of information and a top-down approach to decision-making, are proving inadequate for the effective and responsible integration of AI (Kawakami *et al*., 2024; Taeihagh *et al*., 2021). This hierarchical model, where information ascends for approval and descends for action, is often too slow and rigid to accommodate the dynamic and complex nature of AI and its rapid technological advancements, leading to a significant lag in adaptive response (Chhillar and Aguilera, 2022).

The traditional, linear logic of departmental organisation, characterised by hierarchical structures, procedural workflows, and segmented responsibilities, stands in stark contrast to the demands and dynamics of artificial intelligence (AI) integration. This disjuncture manifests across several critical dimensions, revealing deep structural and epistemological tensions.

AI systems, particularly those employing machine learning (ML), are defined by inherent complexity and uncertainty (Taeihagh *et al*., 2021). Unlike rule-based systems, ML algorithms learn and adapt from vast datasets, producing decisions that evolve dynamically and often exhibit unexpected behaviours. Their internal workings are frequently opaque, making it difficult for human actors to interpret or anticipate outcomes (Chhillar and Aguilera, 2022). This unpredictability renders traditional vertical approval processes slow, sequential, and bureaucratic, ill-suited to manage the emergent risks and rapid changes that characterise AI environments.

The logic embedded within AI systems often diverges from established organisational norms. When AI is adopted, it tends to simplify and codify practices into algorithmic rules, effectively closing" decision-making processes (Gualdi and Cordella, 2021). This algorithmic logic can supersede or conflict with the normative, legal, and administrative principles that traditionally guide public sector operations. The negotiation between AI's rigid regulative regime and the existing institutional, legal, and cultural arrangements within departments creates complex techno-legal-institutional assemblages according to (Gualdi and Cordella, 2021). These assemblages are loosely structured, difficult to integrate, and often resistant to conventional governance mechanisms. As a result, decision-making shifts from context-sensitive human judgment grounded in legal norms to abstract data-driven patterns constructed by algorithms.

The pace of AI development far exceeds the capacity of regulatory and bureaucratic structures to respond. Governments and public institutions, operating within hierarchical and often rigid frameworks, face significant informational disadvantages compared to technology companies and AI developers (Taeihagh *et al*., 2021). This asymmetry leads to regulatory lag, where laws and policies are either too vague, outdated, or insufficiently nuanced to provide effective oversight. The result is a governance vacuum in which accountability and ethical safeguards struggle to keep pace with technological innovation.

The assumption that human oversight can serve as a reliable safeguard within AI-supported decision chains is increasingly challenged. Inserting humans into a linear validation process where AI provides an initial assessment and humans are expected to confirm or adjust it often proves ineffective (Agudo *et al*., 2024). Cognitive biases such as automation bias and anchoring bias lead individuals to over-rely on AI outputs, even when flawed. This creates a "false sense of security," as human oversight becomes quasi-automated, lacking the critical engagement necessary to challenge algorithmic decisions. Moreover, accountability becomes blurred within the human AI assemblage, with responsibility diffused across actors and systems, yet rarely clearly defined (Haesevoets *et al*. 2024).

The integration of AI into public sector decision-making demands a fundamental rethinking of organisational logic. Traditional departmental structures, built for stability and procedural control, are ill-equipped to navigate the fluid, complex, and often ambiguous terrain of AI-driven governance. Addressing this disjuncture requires not only technical adaptation but also institutional innovation, regulatory agility, and a deeper understanding of the cognitive and cultural shift AI introduces.

3.2.1 The Limitations of Traditional Hierarchies (Left Side of the Diagram in Figure 1)

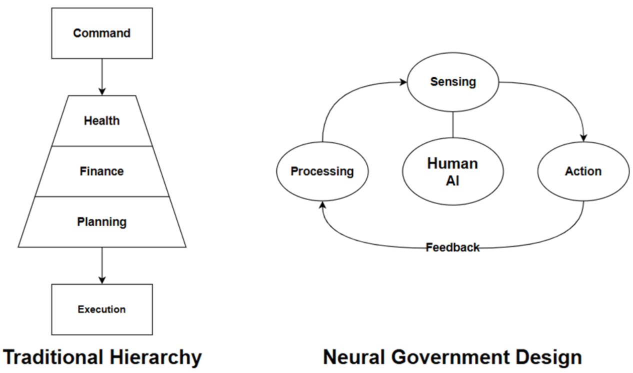

Figure 1: *Traditional Hierarchy vs. Neural Government Design*

Traditional public sector organisations are often structured with a vertical flow of information and a top-down approach to decision-making (Gualdi and Cordella, 2021). This design can lead to slow and rigid processes, hindering adaptive responses (Bovens and Zouridis, 2002).
Information flows up for approval and down for action, a process

that is typically time-consuming. Historically, these structures were designed to accommodate human cognitive limitations by allowing each department to focus on its specific function (e.g., finance, healthcare) (Didin *et al*., 2024).

However, when Information and Communication Technologies (ICTs), including AI, are deployed, they functionally simplify and close organisational practices, encoding them into predefined logical sequences (Kallinikos, 2005).

This process makes the practices embedded in the ICT very difficult to change, creating stable, standardised, but rigid causal interdependencies. The existing normative, legal, and administrative principles that govern public sector decision-making need to be rewritten into the AI's code. This negotiation between the AI's rigid regulative regime and the existing institutional arrangements is complex and time-consuming, creating techno-legal institutional assemblages that are loosely structured and difficult to integrate or untangle (Confalonieri *et al*., 2021).

3.2.2 The Vision of Neural Government Design / Loops (Right Side of the Diagram in Figure 1)

The idea that "information doesn't climb a ladder, it circulates dynamically based on problem relevance, not departmental boundaries" is crucial. This emphasises a shift from command-and-control measures towards more flexible and adaptive approaches (Taeihagh *et al*., 2021).

Such a model allows for continuous feedback, collaboration, and dynamic adjustment, which are essential for navigating the inherent complexity and unpredictability of AI systems. The ability for AI and human agents to exist within feedback loops enables faster learning and cross-domain insight. This moves towards an "adaptive governance" model, which emphasises iterative adjustment and improvement of policies and regulations as new information emerges.

The concept of AI and human agents within feedback loops supports various participatory approaches in AI design and evaluation. This is often conceptualised as human-in-the-loop (HITL) or, more powerfully, human-in-command (HIC).

In an HIC approach, humans leverage AI as a decision aid, but humans retain the upper hand and make the final decisions, with AI providing input and advice rather than making decisions by itself. (Haesevoets *et al*., 2024). This counters the automation bias where humans over-rely on AI suggestions (Agudo *et al*., 2024). Some research suggests that human judgment is more accurate if it is made before receiving erroneous AI support, emphasising the importance of human precedence in the loop (Green, 2022). The lateral and recursive connections between decision nodes enable cross-functional efforts and collaboration across various stakeholders (Mäntymäki *et al*., 2022). This includes incorporating diverse perspectives from frontline workers, legal experts, and impacted communities at earlier stages of AI design and evaluation, fostering a culture of power-sharing and reflexive deliberation. This new organisational logic is vital for achieving responsible AI integration. It ensures that ethical principles like fairness, transparency, and accountability are translated into practicable governance processes.

The diagram in Figure 1 supports the idea that organisations must develop mechanisms and tools to safeguard society from privacy breaches and societal biases that can arise from AI (Chhillar and Aguilera, 2022).

It allows for power-conscious interventions to reshape power relations within agencies and with external institutions and communities (Kawakami *et al*., 2024).

The full end-to-end picture for public sector AI involves a necessary shift from rigid hierarchies to flexible, interconnected Neural Government Design models. This transformation is driven by AI's capabilities and the critical need for responsible deployment, addressing complex challenges through integrated governance frameworks that consider legal, market, normative, and architectural aspects, while continuously balancing various societal trade-offs and ensuring human oversight and accountability in the loop.

3.3 Synthetic Cognition AI as an Organ of Synthetic Cognition

AI is evolving beyond a mere tool, functioning as an organ of synthetic cognition that fundamentally augments human intelligence and, crucially, supports intuition. This redefinition emphasizes AI's role as an integrated partner rather than a subservient instrument (Schmutz *et al*., 2024).

This paradigm shift is driven by concepts such as Augmented Intelligence (AI) or Intelligence Amplification (IA), which are designed to enhance and supplement human intelligence through collaboration (Dave *et al*., 2023). A more advanced form, Artificial Cognition (ACo), distinguishes itself from traditional AI by being brain-inspired and embodied (Damiano and Stano, 2023). Unlike conventional AI's "black box" approach based on data correlations, ACo prioritises proactive knowledge acquisition through dynamic interaction (Sandini *et al*., 2024). It aims to achieve cognitive penetrability, enabling humans and artificial agents to interact based on a shared value system and fostering mutual trust. This positions AI not as an imitator of behaviour, but as a system capable of reproducing underlying organisational mechanisms, reflecting organisational relevance as a deep form of modelling cognition. This collaborative model leads to complementary performance, where human-AI teams surpass the capabilities of either entity acting alone (Berretta *et al*., 2023).

A key aspect of AI acting as an organ of synthetic cognition is its ability to support and enhance human intuition through prospection capabilities, which involves the mental simulation of actions to evaluate their potential future effects, directly supporting informed decision-making by moving beyond trial-and-error (Sandini *et al*., 2024). This "Mental Time Travel" allows cognitive agents to integrate past experiences, present actions, and anticipated future consequences for skilful behaviour. By enabling the mental modelling and evaluation of hypothetical future scenarios, AI systems facilitate proactive reasoning and extend the scope of human intuition.

3.4 Cognitive Architecture of a Loop-Based Government System

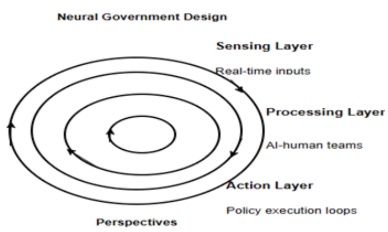

Figure 2. *Three-layer cognitive model of Neural Government Design: sensing real-time signals, processing them through*

AI-human teams, and acting through policy loops informed by feedback.

This model presents the cognitive architecture of an AI-powered public institution, structured into three distinct yet interconnected layers: sensing, processing, and action in seen in Figure 2. These layers are dynamically linked through feedback loops that enable continuous learning, adaptation, and institutional responsiveness.

3.4.1. Sensing Layer

This layer is responsible for capturing raw signals from the external environment, including inputs from citizens, public services, and broader societal conditions. These inputs are multimodal and real-time, encompassing a wide range of data sources. AI technologies are increasingly employed to analyse citizen-generated content, such as public participation data, social media discourse, and digital feedback mechanisms (Mellouli *et al*., 2024). Through natural language processing and sentiment analysis, AI systems can detect arguments, opinions, and emerging concerns. The application of Personal Construct Theory (PCT) and the Repertory Grid technique within information systems research provides methodological tools for understanding how stakeholders conceptualize institutional processes and technologies (Haußmann *et al*., 2012). These approaches offer a nuanced grasp of citizen perspectives, which can inform participatory governance and enhance communication between governments and constituents, including through conversational agents such as chatbots and virtual assistants (Alshahrani *et al*., 2024).

In addition to citizen feedback, AI systems facilitate real-time monitoring of public service usage and environmental conditions. This includes the analysis of operational data from domains such as healthcare, transportation, and emergency response (Mellouli *et al*., 2024). By leveraging both traditional and non-traditional data sources, such as sensor networks, geospatial data, and digital traces AI supports proactive incident detection and service optimisation. The epistemological premise that data becomes a piece of reality underscores the significance of this function: data is not merely representational but constitutive of institutional understanding, offering an objective lens through which societal dynamics are interpreted and acted upon (Alberti, 2019).

To ensure the reliability and relevance of collected data, AI systems perform extensive pre-processing, including noise filtering and pattern recognition. This involves identifying learning sets, selecting appropriate algorithms, and training and evaluating models. The goal is to transform unstructured inputs into structured knowledge that can inform policy decisions and institutional strategies (Mellouli *et al*., 2024). Through these mechanisms, the sensing layer establishes the perceptual foundation of AI-powered governance, enabling institutions to respond to complex social realities with greater precision, agility, and accountability (Alshahrani *et al*., 2024).

3.4.2 Processing Layer

The second layer of the cognitive architecture is characterised by the co-analysis of inputs by human and AI agents, forming what may be described as a synthetic cognition zone. Within this space, AI does not supplant human judgment but rather augments it, enhancing analytical speed, scale, and precision while preserving the irreplaceable qualities of human intuition and contextual understanding. This collaborative dynamic is central to the model's epistemological foundation, wherein cognition is distributed across human and machine actors (Hjaltalin and Sigurdarson, 2024).

Hjaltalin and Sigurdarson (2024) strongly supports the paradigm of human-AI collaboration, often conceptualised through the "centaur" model, in which tasks are optimally performed by a hybrid of human expertise and machine intelligence. AI functions as a form of support intelligence, designed to complement and empower human cognitive, social, and cultural capacities. This approach seeks to harness the creativity and contextual sensitivity of human experts in tandem with the efficiency, accuracy, and scalability of intelligent systems (Henriksen and Blond, 2023). The human-in-the-loop framework is widely advocated, emphasising the importance of human oversight in automated processes and the capacity for intervention when necessary. Such arrangements are particularly vital in domains where ethical judgment, interpretive nuance, and cultural competence are indispensable (Alberti, 2019).

AI systems within this layer can analyse vast and complex datasets with high accuracy, offering structured insights into emerging problems (Prakash *et al*., 2023). Predictive analytics are employed to generate forecasts across a range of public sector applications, including hospital resource planning, crime prevention, and epidemiological modelling. In clinical settings, AI supports medical professionals in diagnosing diseases and formulating treatment plans, demonstrating its utility in augmenting expert decision-making (Mellouli *et al*., 2024).

The outputs generated through this collaborative layer may include scenario models, policy recommendations, and automated decisions, each subject to human review and contextual interpretation. Examples include AI-facilitated risk assessment tools and decision-support systems in public administration (Hjaltalin and Sigurdarson, 2024). However, the deployment of such outputs raises critical concerns regarding bias, fairness, and accountability. The potential for algorithmic systems to reinforce existing social inequities or produce opaque decision-making processes necessitates robust oversight mechanisms. (Alshahrani *et al*., 2024) emphasise that AI systems, while powerful, lack the capacity for human intuition, creativity, and moral reasoning, particularly in complex or ambiguous contexts. These limitations underscore the imperative of maintaining human involvement not only as a safeguard but as an integral component of institutional cognition.

3.4.3. Action Layer

The third layer of the cognitive architecture is responsible for translating analytical insights into concrete policy actions and operational responses. It serves as the executional interface of the system, where decisions informed by prior analysis are enacted within institutional frameworks (Mellouli *et al*., 2024). AI applications in this domain are oriented toward enhancing the efficiency, responsiveness, and quality of public sector operations. These implementations span a wide array of functions, including the optimisation of public service delivery, the streamlining of internal organisational processes, and the pursuit of cost-effective service excellence (Hjaltalin and Sigurdarson, 2024). By automating routine tasks, improving resource allocation, and enabling real time responsiveness, AI contributes to more agile and effective governance.

A defining feature of this layer is its capacity to generate feedback data, thereby establishing recursive loops that connect back to the initial sensing mechanisms. This feedback dynamic is foundational to the system's capacity for learning and adaptation (Kolt *et al*., 2025). As AI-driven actions influence human behaviour and institutional outcomes, the resulting data is reabsorbed into the system, informing subsequent sensing and processing activities. This phenomenon, often referred to as performative prediction, illustrates how AI systems not only interpret reality but also shape it, creating new behavioural patterns and institutional responses that must be continuously monitored and understood.

The feedback mechanism functions as a conduit for organisational learning, enabling the refinement of existing models and the development of new ones. It supports the correction of performance gaps, the updating of institutional memory, and the recalibration of strategic objectives. Effective adaptive governance depends on the ongoing acquisition of up-to-date information regarding AI systems' capabilities, limitations, and societal impacts. This iterative process of assessment and adjustment is essential for maintaining system integrity, preventing the escalation of errors, and ensuring alignment with normative and operational goals. Through this dynamic, the action layer not only executes decisions but also sustains the institution's capacity for reflexivity and resilience.

The Cognitive Architecture of a Loop-Based Government System emphasises organisations as living cognitive systems, not bureaucratic machines. It shows how AI becomes an organ of reasoning, and why looped structures are philosophically and operationally necessary to realise AI's promise in government.

4. NHS – Real Case: From Sequential Triage to AI-Enabled Loops

Current NHS emergency department triage, as described in the Service Design and Delivery: Initial Assessment of Emergency Department Patients (NHS England, 2023), follows a linear, safety-first process. Patients are registered, triaged, and streamed step-by-step, with strict clinical governance. This ensures safety but also builds in multiple handovers, slower escalation, and limited real-time feedback.

Emerging innovations suggest a different approach. Natural language processing (NLP) could analyse patient communications such as NHS 111 transcripts or secure portal messages before they reach the clinical team. AI-driven triage loops would then flag urgent cases, route routine cases to the right service, and learn from every outcome, creating a continuous improvement cycle. Compared to today's process, this model reduces administrative load, accelerates urgent interventions, and continuously adapts to patient demand (see Table 1).

Table 1: Key differences between the current NHS triage framework (NHS England, 2023) and an AI-enabled NLP triage loop (emergent scenario).

Current Process	AI Loop Scenario
Step-by-step escalation through multiple decision points	Real-time analysis and direct routing to the right team
Human-only data capture and prioritisation	NLP parses patient language for urgency and risk
Feedback via audits and governance cycles	Feedback from every case updates the model instantly
Escalation requires handovers	Urgent cases trigger immediate intervention
Capacity based on average demand	Capacity adjusts dynamically to live demand

Why it matters: AI loops could transform NHS triage from a queue-managed process into a responsive, learning network faster for patients, more efficient for clinicians, and better able to adapt in real time.

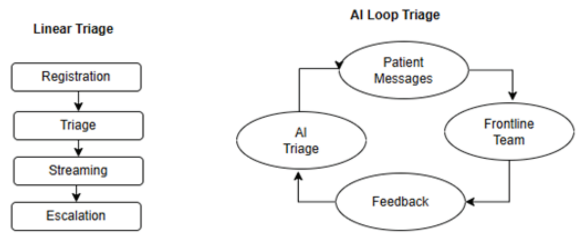

Figure 3. *Comparison of traditional linear triage and AI-enabled feedback loop in NHS emergency care.*

The linear model progresses through sequential stages: registration, triage, streaming, and escalation, each requiring discrete decisions and potential delays. The AI-enabled loop ingests patient messages, performs NLP-based triage, directs cases to the appropriate frontline team, and incorporates outcome feedback to continually refine prioritisation logic.

This transformation is visually captured in Figure 3, which contrasts the traditional linear triage model with an AI-enabled feedback loop. The figure shows how sequential stages, each requiring a separate decision and handover, are replaced by a continuously learning cycle in which patient messages trigger AI triage, frontline action, and outcome-based feedback in near real time. Such a looped structure not only accelerates critical decision-making but also embeds organisational learning into the operational fabric of care delivery.

4.2 GOV.UK Digital Services: Structural Constraints and the Case for Neural Government Design

DSIT (2025) exemplifies both the potential and the limitations of current public sector digital structures. The platform successfully consolidated 1,882 government websites into a single publishing system and is progressing towards a unified One Login for 50 central government services (DSIT, 2025). These achievements demonstrate the value of shared infrastructure in reducing duplication and improving user experience.

However, the State of Digital Government Review reveals that such successes occur in spite of persistent structural barriers. Organisational fragmentation remains a defining feature of the UK public sector: most departments maintain separate technology estates, data standards, and procurement processes, which inhibit interoperability and reusability (DSIT, 2025). In this environment, GOV.UK can centralise access points but cannot, on its own, create the feedback loops, shared data pipelines, or cross-functional governance needed for AI-native service delivery.

The state of digital government review (DSIT, 2025) notes that only 27% of surveyed leaders believe their data infrastructure provides a unified operational view, and 47% of central government services still lack a digital pathway (DSIT, 2025). This misalignment between hierarchical organisational models and the dynamic, iterative nature of AI workflows limits the capacity of GOV.UK to evolve beyond a transactional platform towards a truly adaptive cognitive system.

Applying a Neural Government Design approach to GOV.UK would reconfigure it from a centralised but largely linear access system into a hub within a looped, recursive decision architecture.

Such a model would integrate policy, operational, and digital teams in real time, enable bidirectional data flows across departments, and support hybrid human AI reasoning for more responsive and ethical public services.

5. Intelligence Metrics for Public Service

The integration of Artificial Intelligence (AI) into public sector operations demands a fundamental rethinking of organisational structures, shifting from rigid hierarchies to more adaptive and collaborative loops that prioritise responsible AI integration. While the potential benefits of AI in government, such as enhanced efficiency, improved policymaking, and streamlined service delivery, are widely acknowledged (Van Noordt and Misuraca, 2022), existing evaluation frameworks often fall short in comprehensively capturing the multifaceted impacts of AI on governance and societal values (Berman *et al*., 2024). Current metrics predominantly focus on traditional performance indicators like efficiency or basic accuracy, failing to adequately assess crucial aspects such as human-AI synergy, ethical alignment, and genuine human oversight (Berman *et al*., 2024). This oversight risks the deployment of AI systems that, despite technical proficiency, may undermine public trust, perpetuate biases, or limit human accountability, effectively turning governing with AI into "governing by AI (Van Noordt and Misuraca, 2022).

To truly foster responsible AI integration within public sector "loops," we must move beyond these limited perspectives and introduce metrics that reflect a more holistic understanding of intelligent organisational performance. We propose four new metrics that critically evaluate the success of AI integration, focusing on speed, collaboration, value alignment, and human comprehension, as shown in Figure 4.

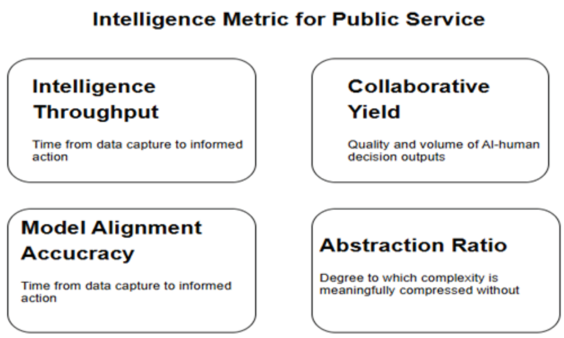

Figure 4. *Intelligence Metrics for Public Service*

5.1 Intelligence Throughput:

Traditional hierarchical structures in the public sector are often characterised by slow information flow and cumbersome decision-making processes, leading to delays in policy responses and service delivery (Bodrick *et al*., 2024; Van Noordt and Misuraca, 2022).

This metric directly assesses the agility and responsiveness of an AI-augmented public sector, measuring how quickly insights derived from data can be translated into actionable decisions and implemented effectively. It moves beyond merely tracking data processing speed to evaluate the full cycle from raw data to a consequential, informed action, which is vital for adapting to rapidly changing societal needs and disruptive forces (Bodrick *et al*., 2024; Doshi, 2025).

5.1.2: Collaborative Yield:

The shift from hierarchies to "loops" necessitates a strong emphasis on human-AI collaboration, recognising that neither humans nor AI alone can achieve optimal outcomes in complex public sector decisions. This metric captures the synergistic value generated by effective human-AI partnerships. It assesses not only the quantity but, crucially, the quality of decisions produced when human intuition, contextual knowledge, and ethical reasoning are combined with AI's computational power, pattern recognition, and data processing capabilities (Kolbjørnsrud, 2024). Measuring collaborative yield ensures that AI is genuinely augmenting human capabilities and that appropriate reliance" is established, allowing humans to distinguish AI advice quality and act upon it effectively (Schemmer *et al*., 2022).

5.1.3 Model Alignment Accuracy:

Unlike the private sector, where AI success might be measured by profit or efficiency alone, public sector AI must inherently align with fundamental public values such as fairness, equality, legality, and accountability (Berman *et al*., 2024; Horvath *et al*. 2023). The sources reveal significant concerns about AI systems failing to meet these critical trustworthiness criteria in practice, often due to inherent biases in data, opacity of algorithms, or lack of proper oversight (Van Noordt and Misuraca, 2022). This metric critically evaluates whether AI recommendations and outcomes are consistent with the ethical and legal principles governing public service delivery. It moves beyond technical accuracy to assess procedural fairness and prevent discriminatory or unreliable results that could erode public trust and legitimacy, addressing the crucial need for a society-in-the-loop approach to AI governance (Horvath *et al*., 2023; Kolbjørnsrud, 2024).

5.1.4 Abstraction Ratio:

For AI integration to be truly responsible and for "loops" to function effectively, human decision makers must understand and trust the AI's output without being overwhelmed by its technical complexity. The challenge of "black box" AI models, particularly neural networks, and the inadequacy of their explanations are significant concerns in the public sector (Berman *et al*., 2024). This metric assesses the AI system's ability to simplify complex data and algorithmic logic into digestible, actionable insights for human users, without compromising accuracy or introducing misleading information. It speaks directly to the principles of interpretability, explainability, intelligibility, and availability. A high abstraction ratio ensures that human judgment and oversight are genuinely informed, enabling caseworkers and managers to meaningfully integrate AI advice into their decisions and fostering a culture of understanding and accountability rather than blind reliance or algorithm aversion (Bodrick *et al*., 2024; Jayakumar *et al*., 2021).

By adopting these four metrics, public sector organisations can undertake a more rigorous and comprehensive assessment of their AI integration efforts. These metrics collectively support the transition from rigid hierarchies to dynamic, adaptive "loops" by prioritising the speed of informed action, the collaborative synergy between humans and AI, unwavering alignment with public values, and the essential human comprehension of AI outputs. This framework offers a critical lens through which to measure not just *what* AI does, but *how well* it supports a responsible, transparent, and effective public service.

5.2 Ethical Considerations on Neural Designs

The integration of AI systems, particularly those built on neural designs, into various sectors has introduced a new paradigm of decision-making, moving from traditional hierarchies to more interactive "loops." However, these advanced systems, often referred to as "black boxes," present significant ethical challenges that necessitate the implementation of sophisticated feedback mechanisms and greater transparency (Baum *et al*., 2022; Chaudhary, 2024). This black box" nature of neural designs is particularly problematic for public sector organisations, as it can lead to a lack of transparency and public distrust in automated decision systems used for public policy, like school admissions or welfare fraud detection (Marian, 2023). It also creates an "accountability gap," making it difficult to assign responsibility when AI systems produce negative or biased outcomes (Baum *et al*., 2022).

To overcome the challenges of opaque neural designs and enable your envisioned looped information flows and "recursive decision structures," robust feedback mechanisms are essential. These are often implemented through Human-in-the-Loop (HITL) systems, which integrate human expertise and oversight into machine learning algorithms and AI processes (Kumar *et al*., 2024; Retzlaff *et al*., 2024). Neural Government Design" offers a forward-thinking vision for integrating AI into public administration. However, for this vision to yield "responsive, ethical, and efficient public services," it inherently demands the sophisticated application of feedback mechanisms (such as HITL) and Explainable AI. These are not merely supplementary tools but are fundamental to addressing the inherent opaqueness of neural designs, fostering public trust, ensuring accountability, and enabling genuine and effective human-AI collaboration in complex decision-making environments (Chaudhary, 2024).

5.3 Strategic Implications in the looped model

In a traditional hierarchical model, public sector leaders often function primarily as approvers, overseeing processes and policies established in static, "one-off" legislative cycles that are slow to change, leading to a "pacing-problem" where regulation lags technological advancements (Reuel and Undheim, 2024). This approach often fails to cultivate a comprehensive understanding of AI's rapidly evolving scope and impact among leaders, hindering strategic success (Maalla, 2021). In contrast, the "looped model" demands that leaders transcend this traditional role, becoming cognitive integrators who actively shape and guide AI integration (Valle-Cruz *et al*., 2024).

Leadership in the era of artificial intelligence (AI) demands a multifaceted approach that integrates strategic foresight, collaborative innovation, and adaptive learning. Central to this transformation is the capacity for strategic and visionary thinking. Leaders must move beyond the superficial adoption of AI technologies and instead formulate long-term visions that embed AI within broader societal frameworks (Valle-Cruz *et al*., 2024). This involves a deep understanding of AI's capabilities and limitations, enabling its effective integration into policymaking, public service delivery, and complex decision-making processes. Such strategic orientation ensures that AI is not merely a technical solution but a catalyst for systemic change.

Equally critical is the role of leadership in fostering collaboration and innovation. In the context of rapid technological disruption, effective leaders must cultivate a culture that embraces experimentation and continuous improvement. They are responsible for inspiring teams to engage with AI constructively, promoting human-AI collaboration that enhances rather than diminishes human agency. This includes providing clear guidance, active support, and the necessary resources for employees to acquire and apply AI-related competencies. By doing so, leaders not only improve organisational efficiency but also empower individuals to navigate and shape the evolving technological landscape.

Finally, continuous learning and adaptability are indispensable traits for leaders operating in the AI-driven Fourth Industrial Revolution (4IR). As AI systems evolve, so too must the leadership mindset, marked by change agility and a willingness to embrace uncertainty. Leaders must be proactive in disseminating knowledge and facilitating skill acquisition across their organisations, preparing the workforce for transformative shifts (Janssen, 2025). Their ability to anticipate technological trends, respond to unforeseen challenges, and manage crises effectively is essential for the successful implementation of AI and the long-term resilience of institutions (Chilunjika *et al*., 2022).

Artificial Intelligence (AI) is no longer a marginal tool for automating discrete tasks; it is an emergent cognitive force that fundamentally reshapes how public institutions sense, decide, and act. As this paper has demonstrated, continuing to retrofit AI into traditional bureaucratic hierarchies built for paperwork, predictability, and slow policy cycles is not only inefficient but increasingly untenable. The rise of AI-native reasoning systems demands a structural shift from linear governance architectures to adaptive, looped systems that reflect the non-linear, feedback rich nature of synthetic cognition (Janssen, 2025; Kolt *et al*., 2025; Mellouli *et al*., 2024).

The persistence of siloed structures, legacy infrastructure, and outdated accountability models impairs the promise of AI in the public sector. High-profile failures like Australia's Robodebt or the Netherlands' SyRI program exemplify how misalignment between technological capability and institutional form can produce systemic harm (Alberti, 2019; O'Connor *et al*., 2024). These cases are not isolated incidents; they are symptoms of a broader epistemological crisis: governing with 21st-century intelligence using 20th-century institutions.

Neural Government Design offers a radical but necessary reconfiguration of public administration. By treating institutions as dynamic cognitive systems capable of sensing environmental inputs, processing through hybrid AI-human reasoning, and acting through iterative policy loops it aligns organizational structure with the functional nature of intelligence (Hjaltalin and Sigurdarson, 2024). Within this framework, AI acts not as an automaton but as an organ of synthetic cognition supporting human intuition, strategic foresight, and ethical judgment (Berretta *et al*., 2023; Sandini *et al*., 2024).

The imperative is not simply to adopt AI but to govern through it intelligently, ethically, and structurally. This requires moving beyond shallow metrics of efficiency toward intelligence-specific indicators such as collaborative yield, intelligence throughput, model alignment accuracy, and abstraction ratio (Berman *et al*., 2024; Bodrick *et al*., 2024). These metrics illuminate whether AI is amplifying human agency or merely displacing it whether it builds trust or obscures responsibility.

The conclusion is clear: governments must stop treating AI as a technical overlay on an outdated machine. Instead, they must recognise it as a structural and epistemic shift one that demands new architectures, new metrics, and a new ethos of co-governance. The choice is stark: cling to hierarchies built for paper, or build loops designed for intelligence. The future of ethical, responsive, and effective public service depends on choosing the latter.

To govern well in an age of intelligence, we must first learn to govern intelligently.

Reference

Adelekan, O. A., Adisa, O., Ilugbusi, B. S., Obi, O. C., Awonuga, K. F., Asuzu, O. F. and Ndubuisi, N. L. (2024) 'Evolving tax compliance in the digital era: A comparative analysis of AI-driven models and blockchain technology in US tax administration', *Computer Science and IT Research Journal*, vol. 5, no. 2, pp. 311–335. doi:10.51594/csitrj.v5i2.759.

Agudo, U., Liberal, K. G., Arrese, M. and Matute, H. (2024) 'The impact of AI errors in a human-in-the-loop process', *Cognitive Research: Principles and Implications*, vol. 9, no. 1, p. 1. doi:10.1186/s41235-023-00529-3.

Alberti, I. (2019) 'The double side of artificial intelligence in the public sector', *Acta Universitatis Sapientiae: Legal Studies*, vol. 8, no. 2, pp. 151–165. doi:10.47745/AUSLEG.2019.8.2.01.

Alshahrani, A., Griva, A., Dennehy, D. and Mäntymäki, M. (2024) 'Artificial intelligence and decision-making in government functions: Opportunities, challenges and future research', *Transforming Government: People, Process and Policy*, vol. 18, no. 4, pp. 678–698. doi:10.1108/TG-06-2024-0131.

Baum, K., Mantel, S., Schmidt, E. and Speith, T. (2022) 'From responsibility to reason-giving explainable artificial intelligence', *Philosophy and Technology*, vol. 35, no. 1, p. 12. doi:10.1007/s13347-022-00510-w.

Berman, A., de Fine Licht, K. and Carlsson, V. (2024) 'Trustworthy AI in the public sector: An empirical analysis of a Swedish labor market decision-support system', *Technology in Society*, vol. 76, p. 102471. doi:10.1016/j.techsoc.2024.102471.

Berretta, S., Tausch, A., Ontrup, G., Gilles, B., Peifer, C. and Kluge, A. (2023) 'Defining human–AI teaming the human-centered way: A scoping review and network analysis', *Frontiers in Artificial Intelligence*, vol. 6, p. 1250725. doi:10.3389/frai.2023.1250725.

Bhima, B., Zahra, A. R. A., Nurtino, T. and Firli, M. Z. (2023) 'Enhancing organizational efficiency through the integration of artificial intelligence in management information systems', *APTISI Transactions on Management*, vol. 7, no. 3, pp. 282–289. doi:10.33050/atm.v7i3.2146.

Bodrick, M., Alqarni, H., Alsuhaim, M. and Almuways, Y. S. (2024) 'Critical appraisal of definitions on intelligence within the organizational context', *Journal of Learning and Development Studies*, vol. 4, no. 2, pp. 12–20. doi:10.32996/jlds.2024.4.2.2.

Bouckaert, G. (2023) 'The neo-Weberian state: From ideal type model to reality?', *Max Weber Studies*, vol. 23, no. 1, pp. 13–59. doi:10.1353/max.2023.0002.

Bovens, M. and Zouridis, S. (2002) 'From street-level to system-level bureaucracies: How information and communication technology is transforming administrative discretion and constitutional control', *Public Administration Review*, vol. 62, no. 2, pp. 174–184. doi:10.1111/0033-3352.00168.

Bullock, J. B. (2019) 'Artificial intelligence, discretion, and bureaucracy', *The American Review of Public Administration*, vol. 49, no. 7, pp. 751–761. doi:10.1177/027507401985612.

Cen, S. H., Ilyas, A., Allen, J., Li, H. and Madry, A. (2024) 'Measuring strategization in recommendation: Users adapt their behavior to shape future content', *arXiv preprint*, arXiv:2405.05596. doi:10.48550/arXiv.2405.05596.

Chaudhary G. (2024) 'Unveiling the black box: Bringing algorithmic transparency to AI', *Masaryk University Journal of Law and Technology*, vol. 18, no. 1, pp. 93–122. doi:10.5817/MUJLT2024-1-4.

Chhillar, D. and Aguilera, R. V. (2022) 'An eye for artificial intelligence: Insights into the governance of artificial intelligence and vision for future research', *Business and Society*, vol. 61, no. 5, pp. 1197–1241. doi:10.1177/00076503221080.

Chilunjika, A., Intauno, K. and Chilunjika, S. R. (2022) 'Artificial intelligence and public sector human resource management in South Africa: Opportunities, challenges and prospects', *SA Journal of Human Resource Management*, vol. 20, p. 1972. doi:10.4102/sajhrm.v20i0.1972.

Confalonieri, R., Coba, L., Wagner, B. and Besold, T. R. (2021) 'A historical perspective of explainable artificial intelligence', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 11, no. 1, p. e1391. doi:10.1002/widm.1391.

Cooper, R. B. and Wolfe, R. A. (2005) 'Information processing model of information technology adaptation: An intra-organizational diffusion perspective', *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, vol. 36, no. 1, pp. 30–48. doi:10.1145/1047070.1047074.

Cristofaro, M. (2022) 'Organizational sensemaking: A systematic review and a co-evolutionary model', *European Management Journal*, vol. 40, no. 3, pp. 393–405. doi:10.1016/j.emj.2021.07.003.

Damiano, L. and Stano, P. (2023) 'Explorative synthetic biology in AI: Criteria of relevance and a taxonomy for synthetic models of living and cognitive processes', *Artificial Life*, vol. 29, no. 3, pp. 367–387. doi:10.1162/artl_a_00411.

Dave, D. M., Mandvikar, S. and Engineer, P. A. (2023) 'Augmented intelligence: Human-AI collaboration in the era of digital transformation', *International Journal of Engineering Applied Sciences and Technology*, vol. 8, no. 6, pp. 24–33. doi:10.33564/IJEAST.2023.v08i06.003.

Didin, D., Akib, H., Haedar, A. W. and Yandra, A. (2024) 'The role of E-government in public services: A bibliometric analysis', *Journal of Contemporary Governance and Public Policy*, vol. 5, no. 2, pp. 111–134. doi:10.46507/jcgpp.v5i2.466.

Doshi, P. (2025) 'From insight to impact: The evolution of data-driven decision making', *Assessment*, vol. 85, p. 92. doi:10.5121/ijaia.2025.16306.

Doğan, K. C. (2020) 'Max Weber'de patrimonyalizm ve bürokrasi kavramları: Antik ve Çin İmparatorluğu üzerine analizler', *Kafkas Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, vol. 11, no. 21, pp. 410–433. doi:10.36543/kauiibfd.2020.019.

DSIT (2025) State of digital government review. *London: Department for Science, Innovation and Technology*.

Dwivedi, Y.K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., et al. (2021) 'Artificial intelligence (AI): multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy', *International Journal of Information Management*, vol. 57, p. 101994. doi:10.1016/j.ijinfomgt.2019.08.002.

NHS England (2023) *Guidance for emergency departments: initial assessment*. Available at: https://www.england.nhs.uk/

guidance-for-emergency-departments-initial-assessment/ (Accessed: 14 December 2025).

Ford, M.R. (2022) 'Making people matter: moving toward a humanity-based public administration', *Administration & Society*, vol. 54, no. 3, pp. 522–539. doi:10.1177/00953997211030213.

Fu, R., Jin, G.Z. and Liu, M. (2022) *Does human–algorithm feedback loop lead to error propagation? Evidence from Zillow's Zestimate*. NBER Working Paper No. 29880. Cambridge, MA: National Bureau of Economic Research. Available at: https://www.nber.org/papers/w29880 (Accessed: 14 December 2025).

Gao, J., Bashan, A., Shekhtman, L. and Havlin, S. (2022) *Introduction to networks of networks*. Bristol: IOP Publishing. doi:10.1088/978-0-7503-1046-8.

González García, C., Núñez-Valdez, E., García-Díaz, V., Pelayo G-Bustelo, C. and Cueva-Lovelle, J.M. (2019) 'A review of artificial intelligence in the Internet of Things', *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 5, no. 4, pp. 9–20. doi:10.9781/ijimai.2018.03.004.

Green, B. (2022) 'The flaws of policies requiring human oversight of government algorithms', *Computer Law & Security Review*, vol. 45, p. 105681. doi:10.1016/j.clsr.2022.105681.

Gualdi, F. and Cordella, A. (2021) 'Artificial intelligence and decision-making: the question of accountability', *Proceedings of the Hawaii International Conference on System Sciences (HICSS 2021)*. doi:10.24251/HICSS.2021.281.

Haesevoets, T., Verschuere, B., Van Severen, R. and Roets, A. (2024) 'How do citizens perceive the use of artificial intelligence in public sector decisions?', *Government Information Quarterly*, vol. 41, no. 1, p. 101906. doi:10.1016/j.giq.2023.101906.

Haußmann, C., Dwivedi, Y.K., Venkitachalam, K. and Williams, M.D. (2012) 'A summary and review of Galbraith's organizational information processing theory', in Dwivedi, Y.K., Wade, M.R. and Schneberger, S.L. (eds.) *Information systems theory*. New York: Springer, pp. 71–93. doi:10.1007/978-1-4419-9707-4_5.

Henriksen, A. and Blond, L. (2023) 'Executive-centered AI? Designing predictive systems for the public sector', *Social Studies of Science*, vol. 53, no. 5, pp. 738–760. doi:10.1177/03063127231163756.

Henriksen, A. and Olesen, F. (2021) 'Experimenting on the enactment of predictive AI: the quest for a future proactive healthcare sector', *STS Encounters*, vol. 12, no. 1. doi:10.7146/stse.v12i1.135404.

Hjaltalin, I.T. and Sigurdarson, H.T. (2024) 'The strategic use of AI in the public sector: a public values analysis of national AI strategies', *Government Information Quarterly*, vol. 41, no. 1, p. 101914. doi:10.1016/j.giq.2024.101914.

Horvath, L., James, O., Banducci, S. and Beduschi, A. (2023) 'Citizens' acceptance of artificial intelligence in public services: evidence from a conjoint experiment about processing permit applications', *Government Information Quarterly*, vol. 40, no. 4, p. 101876. doi:10.1016/j.giq.2023.101876.

Hughes, O.E. (2014) Kamu işletmeciliği ve yönetimi. Ankara: BigBang Yayınları.

Janssen, M. (2025) 'Responsible governance of generative AI: conceptualizing GenAI as complex adaptive systems', *Policy and Society*, vol. 44, no. 1, pp. 38–51. doi:10.1093/polsoc/puae040.

Jayakumar, P., Moore, M.G., Furlough, K.A., Uhler, L.M., Andrawis, J.P., Koenig, K.M., … and Bozic, K.J. (2021) 'Comparison of an artificial intelligence–enabled patient decision aid vs educational material on decision quality, shared decision-making, patient experience, and functional outcomes in adults with knee osteoarthritis: a randomized clinical trial', *JAMA Network Open*, vol. 4, no. 2, e2037107. doi:10.1001/jamanetworkopen.2020.37107.

Kallinikos, J. (2005) 'The order of technology: complexity and control in a connected world', *Information and Organization*, vol. 15, no. 3, pp. 185–202. doi:10.1016/j.infoandorg.2005.02.001.

Kawakami, A., Coston, A., Heidari, H., Holstein, K. and Zhu, H. (2024) 'Studying up public sector AI: how networks of power relations shape agency decisions around AI design and use', *Proceedings of the ACM on Human-Computer Interaction*, vol. 8, CSCW2, pp. 1–24. doi:10.1145/3686989.

Kolbjørnsrud, V. (2024) 'Designing the intelligent organization: six principles for human-AI collaboration', *California Management Review*, vol. 66, no. 2, pp. 44–64. doi:10.1177/00081256231211020.

Kolt, N., Shur-Ofry, M. and Cohen, R. (2025) 'Lessons from complexity theory for AI governance', *arXiv preprint*, arXiv:2502.00012. doi:10.48550/arXiv.2502.00012.

Kumar, S., Datta, S., Singh, V., Datta, D., Singh, S.K. and Sharma, R. (2024) 'Applications, challenges, and future directions of human-in-the-loop learning', *IEEE Access*, vol. 12, pp. 75735–75760. doi:10.1109/ACCESS.2024.3401547.

Kundra, S. and Dwivedi, R. (2023) 'Sensemaking of COVIDian crisis for work and organization', *Philosophy of Management*, vol. 22, no. 1, pp. 129–147. doi:10.1007/s40926-022-00212-5.

Li, H. (2022) 'Application analysis of AI technology in tax collection and administration in China', in *2022 34th Chinese Control and Decision Conference (CCDC)*. doi:10.1109/CCDC55256.2022.10033590.

Ljevo, N. and Šunje, A. (2021) 'The impact of organizational structure on the internationalization strategy', *Economic Review: Journal of Economics and Business*, vol. 19, no. 2, pp. 27–36. doi:10.51558/2303-680X.2021.19.2.27.

Lubana, E.S., Kawaguchi, K., Dick, R.P. and Tanaka, H. (2024) 'A percolation model of emergence: analyzing transformers trained on a formal language', *arXiv preprint*, arXiv:2408.12578. doi:10.48550/arXiv.2408.12578.

Maalla, H.A. (2021) 'Artificial intelligence in public sector: a review for government leaders about AI integration into government administrations', *International Journal of Academic Research in Economics and Management Sciences*, vol. 10, no. 4. doi:10.6007/IJAREMS/v10-i4/11911.

Maciejewski, M. (2017) 'To do more, better, faster and more cheaply: using big data in public administration', *International Review of Administrative Sciences*, vol. 83, no. 1_suppl, pp. 120–135. doi:10.1177/0020852316640058.

Marian, A. (2023) 'Algorithmic transparency and accountability through crowdsourcing: a study of the NYC school admission lottery', *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, pp. 1–12. doi:10.1145/3593013.359400.

Mellouli, S., Janssen, M. and Ojo, A. (2024) 'Introduction to the issue on artificial intelligence in the public sector: risks and benefits of AI for governments', *ACM Journal of Data and Information Quality*, 5, pp. 1–6. doi:10.1145/3636550.

Mäntymäki, M., Minkkinen, M., Birkstedt, T. et al. (2022)

‘Defining organizational AI governance’, *AI Ethics*, 2, pp. 603–609. doi:10.1007/s43681-022-00143-x.

Russell, S.J. and Norvig, P. (2022) *Artificial Intelligence: A Modern Approach*. 3rd edn. Upper Saddle River, NJ: Pearson.

Overman, S. and Schillemans, T. (2022) ‘Toward a public administration theory of felt accountability’, *Public Administration Review*, 82(1), pp. 12–22. https://doi.org/10.1111/puar.13417.

O’Connor, R., Bolton, M., Saeri, A.K., Chan, T. and Pearson, R. (2024) ‘Artificial intelligence and complex sustainability policy problems: translating promise into practice’, *Policy Design and Practice*, 7(3), pp. 308–323. https://doi.org/10.1080/25741292.2024.2348834.

Prakash, A., Jacob, N.E., Merlin, M., Thomas, D.A. and Koshy, S. (2023) ‘Impact of Artificial Intelligence (AI) for decision-making in organisation’, *International Journal of Engineering Technology and Management Sciences*, 7(4), pp. 452–457. https://doi.org/10.46647/ijetms.2023.v07i04.060.

Rességuier, A. and Rodrigues, R. (2020) ‘AI ethics should not remain toothless! A call to bring back the teeth of ethics’, *Big Data & Society*, 7(2), 2053951720942541. https://doi.org/10.1177/2053951720942541.

Retzlaff, C.O., Das, S., Wayllace, C., Mousavi, P., Afshari, M., Yang, T., et al. (2024) ‘Human-in-the-loop reinforcement learning: A survey and position on requirements, challenges, and opportunities’, *Journal of Artificial Intelligence Research*, 79, pp. 359–415. https://doi.org/10.1613/jair.1.15348.

Reuel, A. and Undheim, T.A. (2024) ‘Generative AI needs adaptive governance’, *arXiv preprint*, arXiv:2406.04554. https://doi.org/10.48550/arXiv.2406.04554.

Sager, F. and Rosser, C. (2021) ‘*Weberian bureaucracy*’, in Oxford Research Encyclopedia of Politics. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780190228637.013.166.

Sandini, G., Sciutti, A. and Morasso, P. (2024) ‘Artificial cognition vs. artificial intelligence for next-generation autonomous robotic agents’, *Frontiers in Computational Neuroscience*, 18, 1349408. https://doi.org/10.3389/fncom.2024.1349408.

Schemmer, M., Hemmer, P., Kühl, N., Benz, C. and Satzger, G. (2022) ‘Should I follow AI-based advice? Measuring appropriate reliance in human–AI decision-making’, *arXiv preprint*, arXiv:2204.06916. https://doi.org/10.48550/arXiv.2204.06916.

Schmutz, J.B., Outland, N., Kerstan, S., Georganta, E. and Ulfert, A.-S. (2024) ‘AI-teaming: Redefining collaboration in the digital era’, *Current Opinion in Psychology*, 58, 101837. https://doi.org/10.1016/j.copsyc.2024.101837.

Schwarz, G., Christensen, T. and Zhu, X. (2022) ‘Bounded rationality, satisficing, artificial intelligence, and decision-making in public organizations: The contributions of Herbert Simon’, *Public Administration Review*, 82(5), pp. 902–915. https://doi.org/10.1111/puar.13540.

Selten, F. and Klievink, B. (2024) ‘Organizing public sector AI adoption: Navigating between separation and integration’, *Government Information Quarterly*, 41(1), 101885. https://doi.org/10.1016/j.giq.2023.101885.

Shelmerdine, S., Togher, D., Rickaby, S. and Dean, G. (2024) ‘Artificial intelligence (AI) implementation within the National Health Service (NHS): the South West London AI Working Group experience’, *Clinical Radiology*, 79(9), pp. 665–672. https://doi.org/10.1016/j.crad.2024.05.018.

Sira, M. (2024) ‘Artificial intelligence in knowledge capture and transfer: breaking down organizational silos’, *System Safety: Human–Technical Facility–Environment*, 6, pp. 251-259. https://doi.org/10.2478/czoto-2024-0027.

Stanton, N.A., 2022. Systems-thinking for safety: a short introduction to the theory and practice of systems-thinking: by Simon Bennett, Peter Lang, Oxford, 2019, 172 pp., ISBN: 978-1-78874-377-8 (print). *Ergonomics*, 65(2), p.327. https://doi.org/10.1080/00140139.2021.1992965.

Steele, R., 2025. *Breaking down data silos: navigating the modern data lakehouse landscape*. Arctiq Blog, 17 April. Available at: https://arctiq.com/blog/breaking-down-data-silos-navigating-the-modern-data-lakehouse-landscape (Accessed: 15 December 2025).

Sundar, S.S., 2020. Rise of machine agency: A framework for studying the psychology of human–AI interaction (HAII). *Journal of Computer-Mediated Communication*, 25(1), pp.74–88. https://doi.org/10.1093/jcmc/zmz026.

Taeihagh, A., Ramesh, M. and Howlett, M., 2021. Assessing the regulatory challenges of emerging disruptive technologies. *Regulation & Governance*, 15(4), pp.1009–1019. https://doi.org/10.1111/rego.12392.

Thornton, L., Knowles, B. and Blair, G. (2022) ‘The alchemy of trust: the creative act of designing trustworthy socio-technical systems’, *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’22)*, pp. 1387–1398. https://doi.org/10.1145/3531146.3533196.

Thorstad, D., 2024. *Inquiry under bounds*. Oxford: Oxford University Press.

Valle-Cruz, D., Garcia-Contreras, R. and Muñoz-Chávez, J.P. (2024) ‘Leadership and transformation in the public sector: an empirical exploration of AI adoption and efficiency during the fourth industrial revolution’, *Proceedings of the 25th Annual International Conference on Digital Government Research*, pp. 1–??. https://doi.org/10.1145/3657054.3657146.

Van Noordt, C. and Misuraca, G. (2022) ‘Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union’, *Government Information Quarterly*, 39(3), 101714. https://doi.org/10.1016/j.giq.2022.101714.

Weidener, L., Greilich, K. and Melnykowycz, M. (2024) ‘Adapting Mintzberg’s organizational theory to DeSci: the decentralized science pyramid framework’, *Frontiers in Blockchain*, 7, 1513885. https://doi.org/10.3389/fbloc.2024.1513885.

Hartmann, K. and Wenzelburger, G. (2021) ‘Uncertainty, risk and the use of algorithms in policy decisions: a case study on criminal justice in the USA’, *Policy Sciences*, 54, pp. 269–287. https://doi.org/10.1007/s11077-020-09414-y.

Wright, A.J. (2021) ‘The rise of decentralized autonomous organizations: opportunities and challenges’, *Stanford Journal of Blockchain Law & Policy*, 4(1), pp. 1–27. Available at: https://larc.cardozo.yu.edu/faculty-articles/602 (Accessed: 14 December 2025).

Yildirim, N., Kass, A., Tung, T., Upton, C., Costello, D., Giusti, R. and Meehan, R.O.R. (2022) ‘How experienced designers of enterprise applications engage AI as a design material’, *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI ’22)*, Article 483, pp. 1–13. https://doi.org/10.1145/3491102.3517491.

Yılmaz, V. and Telsaç, C. (2021) ‘Authority and bureaucracy from Weber’s perspective’, *Mehmet Akif Ersoy University Journal of Social Sciences Institute*, 13(34), pp. 42–52.

https://doi.org/10.20875/makusobed.903546.

Zhang, D., Maslej, N., Brynjolfsson, E., Etchemendy, J., Lyons, T., Manyika, J. et al., 2022. *The AI Index 2022 annual report*. arXiv preprint arXiv:2205.03468. https://doi.org/10.48550/arXiv.2205.03468.