

RESEARCH ARTICLE

When AI Delegation Becomes Dependence: Administrative Judgment at Risk

Chengwen Song[†] 

Department of Public Administration, Graduate School, Kookmin University, 77 Jeongneung-ro, Seongbuk-gu, Seoul 02707, Republic of Korea

[†]Corresponding author. Email: wenshao521@kookmin.ac.kr

Abstract

Public organizations can delegate cognitive work to artificial intelligence while losing the capacity to scrutinize or resume that work. When does useful delegation become institutional dependence? This article develops a theory of **institutional AI dependence** as a task–role–workflow configuration combining substantial delegation of core cognition with insufficient task–required review or recovery capacity. It distinguishes two vulnerabilities: inadequate challenge during AI-supported work and inadequate capacity to continue judgment-intensive work when the system is unavailable. Extensive delegation with adequate distributed expertise and tested recovery constitutes managed delegation. Integrating cognitive offloading, automation, organizational learning, and public-administration research, the article extends Song et al.'s (2026) account of epistemic authority to organizational workflows and distinguishes perceived credibility from procedural authority embedded in defaults and override rules. Seven propositions connect practice displacement, challenge behavior, knowledge retention, subsequent delegation, and responsibility–control alignment. They specify how task-relevant practice, diagnostic feedback, and transferable knowledge shape retained capacity over time. The contribution is a task-relative diagnosis of threatened administrative judgment and a set of testable formation mechanisms. The framework guides public organizations in governing current agentic systems and preparing review and recovery arrangements for more capable AI.

Keywords: AI delegation; institutional AI dependence; administrative judgment; public administration; agentic AI; public accountability

1. Introduction

Public organizations can delegate administrative reasoning to AI while retaining responsibility for its consequences. Systems summarize cases, classify applications, recommend actions, and increasingly execute workflow steps. These capabilities can improve throughput and support informed decisions. They also change who examines evidence, recognizes exceptions, and practices the reasoning needed to intervene. The organizational question is when delegation preserves administrative judgment and when it leaves an institution unable to scrutinize or resume the work assigned to AI.

Public decisions apply legal authority, distribute benefits and burdens, interpret exceptions, and require reasons that affected people and oversight bodies can examine. Public organizations remain responsible when AI supplies analysis or performs part of the process. Their institutional capacity includes reviewing consequential outputs, providing reasons, correcting errors, and maintaining



service through disruption. Qualified officials, specialist reviewers, accessible records, and tested fallback arrangements can jointly sustain these functions.

Research on automation explains cognitive offloading and appropriate reliance, while public-administration studies show that responses vary by task and institutional context. Across three experiments, Alon-Barkat and Busuioc (2023) found no automation-bias effect favoring algorithmic over equivalent human-expert advice. Selective adherence to stereotype-consistent advice appeared in one public sample and did not recur among civil servants. Selten et al. (2023) found that street-level bureaucrats trusted recommendations that confirmed professional judgment. A survey of 290 Taiwanese public-school teachers associated ChatGPT dependence with greater work effort through work self-efficacy (Zhang et al., 2025). Deng and Sun's (2025) survey experiment with 190 traffic police linked AI use to lower felt accountability through perceived discretion. These studies identify immediate advice use, effort, and accountability responses. The present theory explains how repeated task allocation can shape later review and recovery capacity.

A recent cognitive-psychological model provides a conceptual bridge. Song et al. (2026) theorize how uncertainty, anthropomorphism, epistemic elevation, and quasi-religious attribution can make AI seem trustworthy beyond verification and normatively authoritative. This article traces how authority attribution becomes embedded in public work roles, training, performance expectations, and accountability. Current agentic systems extend delegation from advice to multi-step execution across records and workflows. As action scope expands, institutions need the capacity to inspect sequences, intervene, and restore service (OECD, 2026a, 2026b). These observable organizational mechanisms provide a basis for studying AGI preparedness while keeping the capabilities of current systems analytically distinct from AGI.

Public-administration theory already connects AI delegation to institutional knowledge loss, impaired scrutiny, and hard-to-reverse dependence (Anastasopoulos, 2026; Wuttke et al., 2026); Wong (2026) locates public judgment in an answerable office. This article refines these arguments through a task-relative diagnosis combining delegation exposure with separate minimum review and recovery capacities. It distinguishes single and combined capacity shortfalls and adequate distributed expertise, then specifies how practice, knowledge retention, and service constraints shape those configurations and subsequent delegation. The contribution is a diagnostic and testable mechanism refinement of established institutional risks.

The central claim is that **institutional AI dependence occurs when substantial delegation of core cognitive work coincides with insufficient task-required review or recovery capacity**. Use frequency describes exposure; trust describes a relationship to system outputs; dependence describes the configuration of delegation and retained capacity. High delegation with adequate review and recovery constitutes managed delegation. Limited delegation with a capacity shortfall identifies a general organizational capability problem.

The argument addresses three questions:

1. Under what task and organizational conditions does AI delegation become institutional dependence?
2. How do practice, authority, and knowledge retention shape the capacity to review and resume delegated administrative work?
3. Which governance arrangements keep AI delegation compatible with effective administrative judgment?

The article makes three contributions. First, it specifies when delegation constitutes institutional dependence by assessing review and recovery against separate task-required capacity floors. Second, it refines existing institutional-dependency arguments through testable links among substitution, authority, practice, knowledge retention, and subsequent delegation. Third, it connects these mechanisms to public-sector arrangements for qualified review, lawful intervention, and service

recovery. Administrative judgment is treated as a capability distributed across roles and workflows, making both its vulnerability and its institutional protection assessable.

Section 2 defines the constructs; Section 3 explains the theory-building approach; Sections 4 and 5 develop the mechanisms and propositions; Section 6 derives institutional design principles; Section 7 sets out boundary conditions; and Section 8 concludes.

2. Conceptual foundations

2.1 *Reliance, overreliance, and institutional dependence*

AI use, reliance, and dependence describe different relationships. **Use** records whether an official or organization employs an AI system. **Reliance** concerns the extent to which a decision-maker uses system outputs when forming judgments or taking actions. Reliance can be appropriate when it corresponds to the system's demonstrated competence and the task's requirements; it can also be too low or too high. J. D. Lee and See (2004) frame trust in automation around appropriate reliance rather than maximal trust; in clinical AI, Benda et al. (2022) adapt this principle, emphasizing the system's purpose, recommendation process, and performance in context. Parasuraman and Riley (1997) distinguish operators' use, misuse, and disuse from designers' and managers' automation abuse, linking these patterns to workload, system reliability, and human-performance consequences.

Overreliance is a decision-level pattern in which a person gives automated advice more weight than evidence or task conditions warrant. Automation bias can involve missing conflicting information; selective adherence involves accepting advice that fits an existing belief or stereotype. These behaviors describe immediate responses, while institutional dependence concerns capability under sustained task allocation.

This article defines **institutional AI dependence** as a task–role–workflow configuration in which substantial delegation of core cognitive work to AI is paired with insufficient organizational capacity to review consequential outputs or continue and recover the work without the system. Delegation exposure is the condition that makes the vulnerability AI-related; challenge and continuity capacities determine whether the organization can govern that delegation. Challenge capacity—the review function during AI-supported work—includes access to relevant evidence, competence to reconstruct outputs, authority to contest them, and time to correct them. Continuity capacity includes qualified backup, usable records, and tested routes for maintaining or restoring service.

The construct explains a specific organizational configuration. Individual overreliance concerns excessive evidentiary weight in a decision; cognitive offloading concerns the transfer of information processing to an external aid; institutional knowledge loss concerns the depletion of organizational knowledge; technological lock-in concerns constraints on changing a system or supplier. Institutional AI dependence identifies whether an organization can govern substantially delegated cognitive work through timely review and recovery. Knowledge loss and switching constraints can contribute to that configuration (Anastasopoulos, 2026), while high delegation with adequate distributed capacity remains managed delegation. The configuration can precede a realized service failure.

Table 1 classifies delegation against task-required challenge and continuity capacity. The standard reflects statutory duties, consequences, discretion, reversibility, and exception demands; capacity may be distributed across trained staff, specialist review, records, and tested fallback procedures.

Table 1. Task-relative configurations of AI delegation and retained capacity

Delegation of core cognitive work	Challenge/review capacity	Continuity/recovery capacity	Organizational configuration
Limited	Adequate	Adequate	Human-led or AI-supported work with retained capability
Substantial	Adequate	Adequate	Managed delegation
Limited	Insufficient	Adequate or insufficient	General organizational capability shortfall
Limited	Adequate	Insufficient	General organizational capability shortfall
Substantial	Insufficient	Adequate	Institutional AI dependence: review pathway
Substantial	Adequate	Insufficient	Institutional AI dependence: recovery pathway
Substantial	Insufficient	Insufficient	Institutional AI dependence: combined pathways

Review and recovery are separate requirements: the first concerns effective intervention during AI-supported work; the second concerns continuation or lawful transfer when AI or a specialist is unavailable. Each is assessed against its own task minimum.

For a benefits eligibility workflow, a program lead working with legal and service-quality owners would map whether AI selects evidence, applies rules, recommends an outcome, or drafts reasons; determine review coverage, notice deadlines, and exception escalation; designate a reviewer with record access, authority, and time to correct; and test how work continues if the system or specialist is unavailable. Delegation is substantial when AI performs or structures a core operation on which the decision depends. Capacity falls below the task floor when required review or recovery cannot be mobilized within statutory and service timelines. Specialist review and tested backup can sustain managed delegation even when frontline staff rarely perform the entire task unaided. Legal and rights-based duties set the minimum floor.

2.2 Administrative judgment and cognitive autonomy

Administrative judgment is the situated cognitive work through which public officials interpret rules, evaluate evidence, recognize exceptions, weigh competing values, and explain decisions. It includes analytical competence, contextual knowledge, professional discretion, and the capacity to revise a conclusion when new information emerges. Public officials exercise this judgment within legal mandates, professional standards, organizational review, and public accountability.

Cognitive autonomy is an official's capacity to evaluate judgments while AI is in use: understand the system's role, compare outputs with evidence, identify uncertainty, and challenge a recommendation. Work units support it through role assignment, records, review time, and authority. **Unaided task performance** is an individual or team's demonstrated performance on a defined task after assistance is removed. **Continuity/recovery capacity** is the unit's available ability to maintain or restore service through qualified backup, accessible records, and tested procedures. Capacity describes available resources and competence; realized service interruption, recovery time, and continuity performance are outcomes observed during an actual event or exercise.

For judgment-intensive tasks, recovery includes reconstructing evidence, applying rules, and giving reasons through qualified roles. This connects continuity capacity to the administrative judgment at stake in delegation. Cognitive offloading can free attention for other work (Risko & Gilbert, 2016). Delegating search or formatting can complement judgment; delegating evidence

appraisal reduces opportunities for independent assessment. Argote and Miron-Spektor (2011) explain learning through task experience, an active context of members, tools, and tasks, and latent organizational conditions. AI changes who practices which task, how feedback is interpreted, and where knowledge is retained. These changes matter where independent assessment, exception handling, or recovery remains required.

2.3 *Epistemic authority and cognitive sacredness*

AI systems can acquire authority through more than formal adoption. A fluent, fast, and apparently consistent system may be perceived as knowledgeable, neutral, or objective. Song et al. (2026) develop the concept of **cognitive sacredness** to describe a threshold at which AI is treated as trustworthy beyond verification and as normatively binding. Their five-stage model links uncertainty, anthropomorphism, epistemic elevation, quasi-religious attribution, and behavioral reliance, while identifying individual, contextual, and technological boundary conditions.

Song et al.'s construct describes human–AI meaning-making and authority attribution. This article extends that account to organizational workflows by tracing how repeated success, standardized procedures, time pressure, and accountability arrangements can embed authority attribution in routine practice. Cognitive sacredness is one possible route to perceived epistemic authority, not a necessary condition for institutional dependence: frequent use can occur without sacralization, and dependence can arise through staffing, workflow, or procurement choices even when individual officials remain skeptical.

2.4 *Public administration, discretion, and accountability*

Information systems can shift administrative discretion from frontline settings toward system design and organizational rules (Bovens & Zouridis, 2002). AI redistributes judgment among frontline officials, managers, data specialists, vendors, and technical systems (Bullock et al., 2020; Young et al., 2019). In public administration, this redistribution takes place within statutory duties to apply rules, consider relevant evidence, explain decisions, and provide review or remedy. The institution therefore needs enough distributed expertise to reconstruct consequential outcomes and act on challenges.

Meaningful accountability requires evidence, time, competence, and authority to disagree. Green (2022) shows that oversight policies often leave these conditions unmet; explanations can also increase trust in incorrect recommendations. Affected people need reasons, usable review, and remedy. Deng and Sun's (2025) intragroup survey experiment with 190 traffic police links AI use to lower felt accountability through perceived discretion. The measure captures officials' self-reported accountability to supervisors under bureaucratic standards. Formal responsibility and affected people's access to remedy require separate assessment.

Anastasopoulos (2026) links default AI decisions and declining oversight expertise to difficult-to-reverse dependency. Wuttke et al. (2026) likewise explain how task automation can erode expertise, weaken scrutiny, and constrain withdrawal; their delegation framework distinguishes assessability, dependency, and contestability. This article operationalizes these concerns through separate task-level capacity floors: review capacity specifies resources for contesting particular decisions, while recovery capacity specifies arrangements for continuing, transferring, or reducing delegated work. A unit with strong reviewers but no fallback differs from one with tested recovery but ineffective challenge. P1/P5/P6 specify practice and knowledge-retention relationships; P4 specifies conditions for subsequent delegation. These distinctions refine the diagnosis and predictions of existing institutional-dependency arguments.

2.5 *Agentic AI and the scope of delegated action*

The OECD review distinguishes autonomy—the capacity to act without direct human involvement—from broader agency, which includes goal formulation, long-horizon reasoning, and adaptation; autonomous operation alone does not establish these fuller agentic capacities (OECD, 2026a). Because definitions vary, this article uses a task-focused definition of agentic AI: systems that can decompose goals, use tools or interact with digital workflows, and pursue multi-step work over time with limited human supervision. Their action autonomy ranges from recommendation-only systems, through human-approved or human-monitored execution, to bounded action without step-by-step involvement. Government applications can retrieve records, route cases, update files, or initiate transactions. Review therefore extends from final outputs to action sequences, permissions, and state changes (OECD, 2026a, 2026b).

The theory applies across this range when AI takes on core cognitive operations. Greater action autonomy increases the importance of reconstructable records, intervention authority, and tested recovery because an error can propagate across workflow steps before a human reviews the final outcome. The AGI implication is institutional preparedness: agencies need review and recovery arrangements that remain usable as task scope and autonomy expand. The propositions concern organizational mechanisms observable in current AI-supported work; the extension to AGI concerns the scale and reach of delegated action.

3. Theory-building approach

The article uses an integrative conceptual design linking cognitive offloading, automation, human–AI decision-making, organizational learning, administrative discretion, and accountability (Jaakkola, 2020). Analysis began with the 21 full-text works already collected and cited for the project: one methodological and 20 substantive sources. Full-text review retained works that defined a construct, explained a mechanism, specified a public-sector condition, or informed a safeguard. Bibliographic and claim matching was completed on 19 October 2026. Three targeted additions addressed gaps exposed during revision: Argote and Miron-Spektor (2011) on organizational knowledge retention, Deng and Sun (2025) on felt accountability, and OECD (2026a) on agentic capability. These were located through topic- and title-based searches of publisher and official-agency sites through that date. The source set is purposive. Its recoverable record comprises the 24 full texts, the source audit, and the item-level map in Supplementary Appendix A; the original discovery and exclusion history of the core collection was not retained. Earlier project database exports therefore supply no screening denominator for this conceptual article.

The synthesis distinguishes delegation exposure from retained capacities; separates offloading, epistemic authority, and procedural authority; traces practice and knowledge retention across roles; and derives testable relationships. Divergent findings about advice use, effort, and accountability inform boundary conditions. Table 2 summarizes the source streams; Appendix A records each work's theoretical role, proposition link, study setting, measured outcome, and transfer boundary where applicable.

Empirical studies support claims about their measured populations and outcomes. Evidence from industrial and dynamic automation settings motivates practice and recovery mechanisms; it does not establish long-term skill change among public officials using generative or agentic AI. Survey experiments and observational studies inform immediate reliance, discretion, effort, or accountability responses. Conceptual and policy sources inform public-sector conditions and institutional design. The article therefore presents cumulative capability change as a theoretical mechanism for future testing.

Table 2. Included sources, theoretical roles, and evidence boundaries

Source stream and included sources	Contribution to the framework	Mechanisms or propositions informed	Evidence boundary
Conceptual research design: Jaakkola (2020)	Grounds the integrative conceptual approach and proposition development	Research design; no substantive proposition	Methodological guidance; not evidence for the proposed organizational mechanisms
Cognitive offloading and automation: Risko and Gilbert (2016); Bainbridge (1983); Endsley and Kiris (1995)	Define offloading and motivate links among task allocation, practice, monitoring, and performance after automation	Practice displacement, monitoring, and recovery; P1 and P5	Review and automation studies motivate mechanism analogies; they do not directly estimate long-term AI effects in public agencies
Human–AI reliance and cognitive authority: Benda et al. (2022); J. D. Lee and See (2004); Parasuraman and Riley (1997); Alon-Barkat and Busuioc (2023); Selten et al. (2023); H.-P. Lee et al. (2025); Song et al. (2026)	Distinguish appropriate reliance, over-reliance, advice use, reported cognitive effort, and perceived epistemic authority	Epistemic authority and challenge behavior; P2; task allocation and workflow reliance inform P3	Findings concern particular systems, tasks, samples, and immediate or self-reported outcomes; Song et al. provide a conceptual account of authority attribution
Public administration and discretion: Bovens and Zouridis (2002); Bullock et al. (2020); Young et al. (2019); Zhang et al. (2025)	Locate cognitive work in discretion, role design, and public-service work	Workflow authority and role allocation; P3; institutional context for P7	Studies concern specified settings and outcomes; P4's subsequent-delegation relationship is derived in this article
Accountability and public value: Bovens (2007); Green (2022); Anastasopoulos (2026); Wong (2026); Wuttke et al. (2026)	Specify answerability, review, correction, contestability, and responsibility–control relations	Challenge and remedy conditions; responsibility–control alignment; P7	Conceptual and empirical contributions inform institutional pathways; their constructs remain distinct from dependence
Government AI context: OECD (2026b)	Informs public-sector transformation and institutional readiness	Institutional scope and safeguards; context rather than direct evidence for P1–P7	Policy analysis describes governance needs; it does not test the propositions
Targeted full-text additions: Argote and Miron-Spektor (2011); Deng and Sun (2025); OECD (2026a)	Specify cross-role knowledge retention and transfer, an immediate street-level accountability pathway, and agentic-AI capability distinctions	Organizational knowledge retention; P6; accountability context for P7; agentic-AI scope conditions	Argote and Miron-Spektor provide a cross-level learning mechanism; Deng and Sun measure self-reported felt accountability and discretion in traffic-police scenarios; OECD (2026a) defines agentic capability categories. None tests the article's longitudinal propositions.

P4 is a deductive feedback proposition combining continuing service obligations, reduced retained capacity, perceived AI viability, and limited alternatives. The included sources inform these inputs; none directly tests their complete feedback relationship. The framework covers decision-support and workflow systems that perform or structure public administrative cognition.

4. From AI delegation to institutional dependence

The model separates a task-allocation decision from the cognitive and procedural processes that follow it. **Substitution exposure** describes the share of core reasoning assigned to AI. **Cognitive offloading** describes how an official uses the system to reduce internal processing. Epistemic authority concerns the credibility assigned to the output; procedural authority concerns the position the output receives in the workflow. These processes can occur together, but they have distinct causes and indicators. Repeated substitution can reduce task-relevant practice; review, staffing, performance rules, and accountability arrangements shape whether lower practice becomes an institutional vulnerability.

4.1 Cognitive offloading

AI can lower the effort required to search, summarize, compare, classify, and draft. Under heavy caseloads or limited time, delegation can free attention for complex cases. Risko and Gilbert's (2016) account of cognitive offloading explains why people use external aids when internal processing is demanding. In AI-supported work, the same process can reduce engagement with underlying information when a system supplies both an answer and a rationale.

In public administration, the relevant distinction is between **offloading operations** and **offloading judgment**. An official may use AI to retrieve documents while independently deciding which evidence matters. Alternatively, the system may select evidence, assign relevance, recommend an outcome, and generate the reasons that make the outcome appear defensible. The greater the bundle of cognitive tasks delegated, the fewer opportunities the official has to form an independent assessment.

4.2 Epistemic authority transfer

Offloading is more consequential when the AI output is treated as an authoritative account of the case. An answer's fluency, speed, quantitative appearance, and integration into official systems can make it seem more complete or objective than it is. Song et al.'s (2026) cognitive-sacredness model helps explain how perceived authority can exceed verification. J. D. Lee and See (2004) clarify a complementary process: trust guides reliance especially when system complexity and unexpected conditions make complete understanding impractical. In organizations, this reliance becomes consequential when workflows make an AI output the default basis for action and independent checking costly. Managers may reinforce this pattern by standardizing system recommendations, while staff who rarely observe system failure or must justify disagreement face fewer incentives to question them.

Epistemic authority changes the perceived evidentiary weight of an output. Officials may use it as the baseline and search for reasons to deviate, creating anchoring and reducing attention to

contradictory evidence. Public-sector evidence shows context-dependent responses: Alon-Barkat and Busuioc's (2023) Study 2 found selective adherence to stereotype-consistent advice among members of the public, while Study 3 did not reproduce this pattern among civil servants. The contrast indicates that advice use varies by sample and setting.

Epistemic authority transfer is therefore expected to be stronger when outputs are presented as precise, explanations are difficult to inspect, workload makes independent checking costly, and organizational norms frame system compliance as prudent. It should be weaker when staff can compare the output with case evidence, see system uncertainty, and challenge recommendations without penalty.

Procedural authority changes the action environment. A recommendation acquires procedural standing when it appears as the default in a case system, is embedded in a standard operating procedure, or requires extra justification to override. Officials may remain skeptical of the model and still follow that default. Epistemic authority should be measured through perceived credibility and evidentiary weight; procedural authority should be measured through system defaults, permissions, approval steps, and override costs. The distinction separates belief-driven challenge from workflow-driven compliance.

Challenge behavior is informative when an error, conflict, or exception creates a task-relevant opportunity for independent review. Correct acceptance of sound advice is calibrated reliance. Repeatedly missed review opportunities can reduce corrective feedback and learning; any subsequent loss of organizational capacity operates through practice, knowledge retention, and access to qualified alternatives. The frequency of challenge alone cannot establish that learning path.

4.3 Practice displacement and competence erosion

Professional competence develops through task experience, feedback, and opportunities to apply knowledge (Argote & Miron-Spektor, 2011). When AI performs routine analytical tasks, employees may gain time and access to better information. When it performs tasks through which officials acquire and maintain judgment, work can shift from active assessment to monitoring and exception handling. Bainbridge's (1983) analysis of industrial automation describes a related pattern: operators may retain responsibility for abnormal conditions while routine control reduces practice opportunities. This mechanism directs attention to task exposure, feedback, and readiness to intervene.

In administrative work, practice displacement may affect the ability to interpret policy, identify atypical evidence, recognize a false premise, explain a decision, or respond to a system outage. Training should therefore cover the underlying administrative task as well as tool operation. Dependence may also grow when agencies reduce staffing or remove manual processes on the assumption that AI performance will remain continuously available.

Automation research links reduced engagement with situation-awareness and performance problems in particular control settings (Bainbridge, 1983; Endsley & Kiris, 1995). These studies motivate the possibility that officials who rarely practice a task may be slower to identify missing evidence, policy changes, or anomalous cases; they do not directly measure those administrative outcomes. A survey of knowledge workers associates greater confidence in generative AI with lower self-reported effort on some critical-thinking tasks (H.-P. Lee et al., 2025), while public-administration studies measure context-specific advice use (Alon-Barkat & Busuioc, 2023; Selten et al., 2023). Together, these sources motivate the proposed trajectory from repeated substitution to reduced practice and task-specific performance. Longitudinal competence loss among public officials remains a theoretical prediction for future testing.

Practice displacement depends on task match and feedback. Practice in a related activity may not preserve the judgment needed for a particular decision, while rehearsal that exposes reasoning and returns diagnostic feedback can sustain learning. The evidence also separates work effort from unaided competence: Zhang et al. (2025) associate ChatGPT dependence with greater reported

effort through work self-efficacy, whereas H.-P. Lee et al. (2025) report lower perceived effort on some critical-thinking activities. These outcomes measure different processes. The framework predicts task-specific change in independent performance and organization-level recovery; effort and confidence remain separate outcomes.

4.4 *Accountability asymmetry as organizational feedback*

Organizational accountability shapes whether cognitive dependence can be challenged and corrected. Public officials often remain responsible for decisions even when AI systems structure the evidence, recommend an outcome, or execute workflow steps. If an official cannot explain why the system's recommendation was accepted, identify what information it used, or independently evaluate a counterfactual, formal responsibility can outlast practical control. The result is **accountability asymmetry**: an organization attributes responsibility to human actors while making it difficult for them to exercise the cognitive control that responsibility presumes.

This asymmetry can reinforce dependence. Officials may follow a system they cannot confidently challenge, while managers may treat a human signature as evidence of oversight. A signature can record a procedural step without showing that the reviewer assessed the evidence. Procurement choices, system integration, staff specialization, and performance systems may also raise the cost of switching or returning to manual work, turning individual reliance into an organizational routine (Anastasopoulos, 2026).

Accountability quality depends on whether the responsible office-holder can access relevant evidence, understand the system's role, explain the decision in administrative terms, and correct or reverse it. For Bovens (2007), accountability links an actor to a forum before which conduct is explained, questioned, judged, and subject to possible consequences. Logs and explanations support that relationship when an office-holder and oversight forum can interpret, question, and act on them (Bovens, 2007; Green, 2022).

Accountability arrangements shape review, workflow authority, and later task allocation. Record access, review time, and intervention authority support challenge capacity; defaults and override costs shape challenge behavior; correction and audit feedback can preserve task knowledge. Affected people need usable routes to contest decisions and obtain remedies. These arrangements are institutional conditions. The quality of reasons, correction, and remedy constitutes downstream accountability outcomes.

5. Framework and theoretical propositions

5.1 *Construct and unit of analysis*

The unit of analysis is a **task-role-workflow configuration**. Institutional AI dependence combines substantial cognitive delegation with a shortfall in **challenge/review capacity** during system use or **continuity/recovery capacity** during interruption or transfer. The two capacity dimensions have distinct indicators and governance responses. An agency can retain strong specialist review yet lack a usable recovery route, or maintain service while lacking the ability to challenge consequential outputs.

Challenge capacity concerns the resources an organization can mobilize within the task's legal and service timelines: relevant evidence, qualified expertise, time, and effective authority. Challenge behavior consists of observed checking, questioning, escalation, or correction. Unaided performance and behavior on error or exception cases inform capacity assessment; the assessment asks whether required functions can actually be performed. Specialist expertise counts when it is accessible in time and empowered to intervene. Actual disruption and recovery performance provide further observations of continuity capacity.

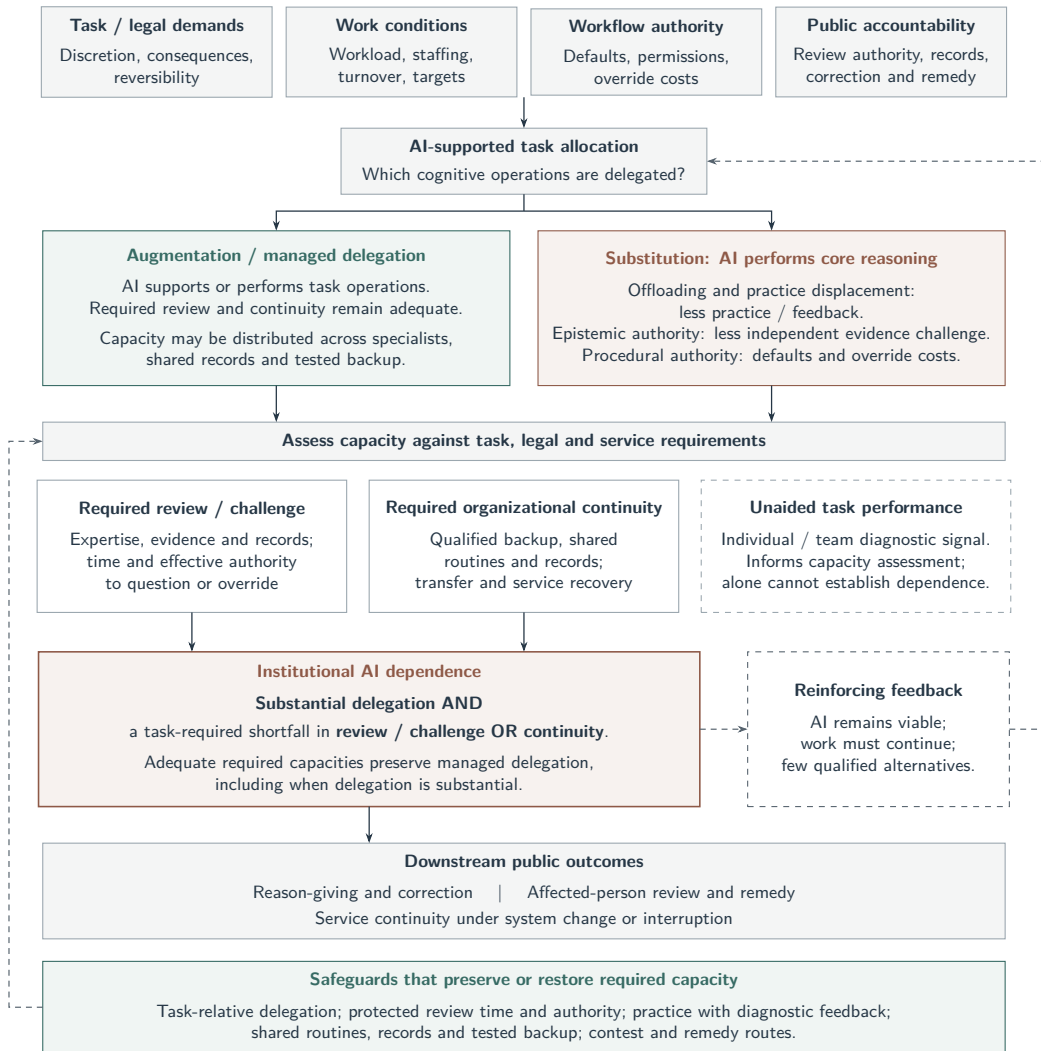


Figure 1. Task allocation, retained capacity, and institutional AI dependence. The diagram combines proposed mechanisms with a diagnostic process. Assessment connections indicate evaluation against task requirements. Classification requires substantial delegation and a required review or recovery shortfall. Feedback begins with an independently observed decline in task performance or recovery capacity under P4's service, viability, and alternative-capacity conditions. Top-row accountability items are institutional arrangements; downstream items are public outcomes evaluated through the specified mechanisms, including P7.

5.2 Managed delegation and institutional dependence

The framework links task allocation to institutional capability through three connected pathways. First, repeated substitution for core reasoning can displace task-matched practice and reduce unaided performance. Second, perceived epistemic authority and workflow-based procedural authority can weaken challenge while the system is active. Third, reduced task and recovery capacity can make further delegation attractive when work must continue and qualified alternatives are limited. Review rights, task-relevant practice, feedback, staffing, and fallback design shape each link.

An **augmentation pathway** emerges when AI handles retrieval, formatting, or comparison while officials retain problem framing, evidence appraisal, and reason-giving. **Managed delegation**

can also include substantial automation when the organization retains task-appropriate review, records, and continuity capacity. An **institutional dependence pathway** emerges when core reasoning is delegated and the organization cannot mobilize sufficient challenge or continuity capacity. These patterns can coexist across tasks and roles within one agency. Both can produce short-term throughput gains; the distinction concerns whether public organizations can still examine consequential outputs and sustain lawful service when AI is challenged or unavailable.

Proposition 1: Substitution and task-specific practice

P1. Holding task demands and compensating learning opportunities constant, repeated AI substitution for core, task-specific reasoning predicts lower subsequent unaided task performance.

P1 states a baseline longitudinal relationship for tasks requiring independent judgment, exception recognition, or resumption. P5 specifies how variation in task-matched practice and diagnostic feedback changes that relationship.

Proposition 2: Epistemic authority and challenge behavior

P2. When a task calls for independent verification, stronger organizational or individual attribution of epistemic authority to AI outputs predicts less search for disconfirming evidence and fewer evidence-based challenges, particularly under time pressure and when system performance is difficult to inspect.

The outcome is challenge behavior when independent review is required and an error, conflicting evidence, or exception creates an opportunity to challenge. Perceived credibility and evidentiary weight are measured separately from that behavior.

Proposition 3: Procedural authority and challenge behavior

P3. When workflow defaults, permissions, approval sequences, or override costs give AI outputs procedural priority, officials are less likely to conduct required independent checks or raise documented exceptions, including when their private assessment of system credibility remains cautious.

This proposition separates workflow position from individual belief. Its indicators include default settings, permissions, mandatory approval steps, required override explanations, and the time needed to depart from a recommendation.

Proposition 4: Capability loss and subsequent delegation

P4. When affected work must continue, managers regard an available AI system as a viable way to maintain output, and qualified short-term alternatives are limited, a decline in task-relevant performance or recovery capacity increases the likelihood of further delegation, especially under workload or throughput pressure.

P4 is deduced from the interaction of service obligations, retained capacity, and available alternatives. AI viability judgments and resource constraints must be measured before or independently of subsequent delegation. A decline can also prompt reduced AI scope, qualified review, training, or backup. When the stated conditions hold and delegation does not increase, the observation counts against P4.

Proposition 5: Task-relevant practice as a safeguard

P5. Task-relevant practice with diagnostic feedback weakens the relationship between repeated AI substitution and declines in unaided task performance.

Practice, review, and backup can be distributed across roles. High-consequence discretionary work may require an independent first assessment or trained specialist review; routine and reversible work may rely on sampling, exception review, and periodic manual exercises. Each arrangement

must meet the task's legal and service requirements. Experience also requires interpretation: rare or ambiguous outcomes can produce inappropriate organizational lessons, which makes diagnostic feedback and shared review important to knowledge creation (Argote & Miron-Spektor, 2011).

Proposition 6: Organizational retention of task knowledge

P6. Task-relevant experience strengthens unit-level challenge or recovery capacity when diagnostic feedback validates learning and the resulting knowledge is retained and transferable through shared routines, accessible records, and cross-role expertise networks.

Argote and Miron-Spektor (2011) distinguish member learning from knowledge embedded in repositories and transfer channels others can access. P6 tests this cross-level translation. Feedback accuracy, record accessibility, and transfer of task knowledge should be measured independently of performance on new review cases or recovery exercises, allowing effective learning arrangements to be evaluated against later capacity.

Proposition 7: Responsibility-control alignment and public accountability

P7. In consequential public decisions, a larger gap between an office-holder's formal responsibility and the practical ability to reconstruct, challenge, and correct AI-supported action is associated with weaker reason-giving and slower correction for affected people and oversight forums.

P7 treats reason-giving and correction as downstream accountability outcomes; institutional AI dependence is classified from delegation exposure and task-required challenge or continuity capacity. The public-sector relevance follows from duties to explain administrative action, respond to review, and provide remedy.

Table 3. Propositions and candidate empirical indicators

Prop.	Core relationship	Measurement roles and candidate indicators
P1	Substitution → lower later unaided performance, other learning opportunities held constant	Predictor: share of core reasoning delegated repeatedly. Outcome: task-matched and novel-case accuracy after assistance is removed.
P2	Epistemic authority → reduced required challenge	Predictor: credibility and evidentiary weight. Conditions: review requirement and error, conflict, or exception. Outcomes: evidence search and justified challenge; time pressure and inspectability are moderators.
P3	Procedural authority → reduced required checks or documented exceptions	Predictors: defaults, permissions, approval order, override time and cost. Outcomes: checks and correction on seeded-error cases. Credibility beliefs are measured separately.
P4	Capacity decline → further delegation under specified constraints	Predictor: independently measured capacity decline. Prior conditions: continued service need, perceived AI viability, limited qualified alternatives. Outcome: later delegation scope. Competing responses: restriction, training, review, or backup.
P5	Practice and feedback moderate P1	Moderator: amount of task-matched practice and diagnostic feedback. Outcome: later unaided performance; substitution exposure is measured separately.
P6	Retained, transferable learning → unit capacity	Inputs: diagnostic feedback, record access, and demonstrated cross-role knowledge transfer. Outcomes: unfamiliar-case review and tested recovery performance.
P7	Responsibility-control gap → weaker accountability outcomes	Predictor: formal duties relative to reconstruction and intervention authority. Outcomes: reason quality, correction time, and usable remedy.

Table 3 separates predictors, conditions, and outcomes. Capacity tests ask whether competent review or recovery can be mobilized within required timelines; behavior is observed where review

is warranted. Studies should report these measures alongside delegation exposure before classifying institutional AI dependence.

5.3 Cross-level logic

Argote and Miron-Spektor (2011) locate organizational knowledge in members, tools, routines, and their networks. Learning becomes organizational when others can access that knowledge. Repeated AI substitution changes these relationships as tools execute more core reasoning and store more task knowledge. Practice and feedback losses become institutional vulnerabilities when shared repositories and transfer channels are weak. Rotation with handoffs, cross-training, specialist review, records, and tested backup can distribute capacity; staffing cuts, narrow specialization, and throughput incentives can concentrate it in a tool or a few specialists.

Dependence may concentrate in a role, case type, or transition point: a unit can process routine cases effectively while losing capacity to manage exceptions or train replacements. Institutional capacity may be distributed; every frontline official need not perform every task unaided. Agencies should map which roles retain task knowledge, who can challenge AI outputs, what records support reconstruction, and whether the unit can maintain service when a system or specialist is unavailable.

This cross-level sequence clarifies the framework's public-administration contribution. Individual challenge and task performance describe frontline capabilities; work-unit records, role allocation, and backup describe organizational continuity. Statutory duties and affected people's rights to reasons, review, and remedy determine the task-relative floor for those capacities. The model therefore treats individual skill and organizational resilience as related but non-equivalent outcomes.

6. Governing delegation while retaining administrative judgment

Managing institutional AI dependence requires matching delegated work to task-appropriate review and continuity capacity. The design objective is an organization that uses AI where it adds value and can still examine, correct, or resume consequential work.

6.1 Match delegation to task and consequence

Agencies should distinguish information retrieval and formatting from evidence evaluation, exception handling, and consequential judgment. Delegation can be broader for routine, reversible tasks with clear criteria when a functioning review and fallback route remains available. Decisions involving statutory discretion, disputed evidence, exceptions, or rights require stronger independent assessment, reasons, and accessible reconsideration. The benefits-review example in Section 2.1 illustrates how agencies can map delegated operations, required reviewer authority, timelines, and fallback coverage. Statutory and rights-based duties establish the minimum floor; throughput gains cannot substitute for those requirements. Human discretion can also produce inconsistency or bias, so the design standard is task-appropriate review and remedy, not universal manual performance. A task's risk profile should consider immediate error costs and whether the organization can challenge and recover the work after prolonged delegation.

6.2 Preserve independent practice

Training and work design should preserve the task knowledge required for meaningful review and service recovery. Agencies can use periodic manual case exercises, planned rotations with handoffs, simulations of system outages, and review of cases where AI advice conflicts with other evidence. Because individual learning reaches the organizational level when knowledge is embedded in shared repositories and transfer routines, agencies should pair exercises with accessible records, cross-role expertise, and tested fallback procedures (Argote & Miron-Spektor, 2011). The required practice level varies with task consequence, discretion, and fallback design.

Public personnel systems can assign ownership for practice and backup. Program managers identify the tasks and roles requiring independent review; supervisors allocate review time and exercises; HR embeds task knowledge in onboarding and professional development; procurement and technical teams preserve usable records and workable fallback access. Supervisors assess evidence review, escalation, correction, and reasoned disagreement alongside service quality. When AI reduces routine workload, managers can reserve part of the capacity for case learning, feedback, and recovery exercises. These arrangements translate task-level safeguards into staffing and performance-management decisions (Argote & Miron-Spektor, 2011; Endsley & Kiris, 1995; OECD, 2026b).

6.3 Design challenge into the workflow

A meaningful review role gives officials the evidence, time, and authority needed to assess and contest AI-supported action. Interfaces can surface uncertainty, missing information, alternatives, and system limits. For consequential decisions, an initial human assessment can precede the AI recommendation. Workflow defaults and permissions should make a reasoned override operationally possible, and managers should evaluate the grounds for challenge rather than treat deviation itself as noncompliance.

6.4 Align accountability with cognitive control

Institutions should equip legally responsible offices and roles with the information, competence, and authority needed to discharge their duties. Internal assignments must follow the governing legal framework. Accountability maps should identify who defines delegated tasks, authorizes access, reviews outputs, pauses execution, and restores service. Review policies should specify evidence access, time, intervention powers, and affected people's routes to remedy. Supervisors, program leads, procurement and technical teams, auditors, ombuds offices, and courts need records and intervention routes appropriate to their statutory remits. These arrangements support the answerable office required for public judgment (Bovens, 2007; Wong, 2026).

6.5 Evaluate capability alongside throughput

AI evaluations should compare service quality during use, available review and recovery capacity, and realized performance during challenge or interruption. Measures include error detection, exception handling, reason quality, remedy access, qualified backup, and recovery time. A dependence classification identifies a vulnerability; deployment choices require comparing feasible options, including existing service provision, strengthened specialist backup, and narrower delegation. AI can improve an initially understaffed service while leaving recovery vulnerability. The preferred arrangement must satisfy legal duties and improve service against relevant alternatives.

6.6 Prepare for more capable and agentic systems

Institutional readiness includes the ability to define task scope, limit permissions, inspect actions, pause execution, and restore service. These capacities become more consequential when systems execute multi-step work across records and workflows. OECD guidance on government transformation emphasizes bounded permissions, observable actions, retained accountability, and routes to contest decisions (OECD, 2026b); its conceptual analysis treats agentic capability as a range rather than a single autonomy threshold (OECD, 2026a). These principles apply to current systems and provide an institutional basis for AGI preparedness.

7. Boundary conditions and future research

The relationships vary with task complexity, legal discretion, reversibility, workload, output inspectability, and practice opportunities. Substitution for complex or exception-sensitive reasoning

carries greater dependence risk when review and recovery capacity is weak. Routine, reversible tasks can support broader delegation when records, sampling, escalation, and fallback are adequate. Human review should also account for the risk that discretionary judgment can produce inconsistent or biased treatment.

The task–role–workflow combination remains the unit of analysis. The same system can augment retrieval in one role, substitute for evidence appraisal in another, and act on case records in a third. Dependence is classified only when substantial delegation coincides with a task–relevant shortfall in challenge or continuity capacity. The required capacity may be held by a distributed team, specialist office, or tested operational fallback.

Staffing reductions, turnover, outsourcing, procurement choices, policy complexity, and changes in case mix can also reduce retained capability or increase delegation. The construct classifies a present task–role–workflow configuration. Attribution to AI requires measuring baseline capacity and subsequent change alongside task allocation, practice, workforce changes, restructuring, and administrative demands. These factors supply rival formation pathways and boundary conditions.

Task experiments can compare AI–first, human–first, and collaborative workflows on accuracy, calibration, evidence review, and later unaided performance. Longitudinal studies can follow tasks before and during deployment, then observe review and recovery during outages, policy changes, and contested cases. Organizational designs should measure staffing, turnover, task composition, procurement, training, targets, and legal rules alongside delegation. Logs can identify evidence searches, edits, escalation, and correction; interviews and documents can establish how expertise and remedy are distributed. Across designs, delegation, behavior, review capacity, recovery capacity, and realized outcomes require separate measures and baseline observations.

The theory focuses on review and continuity capacity. Discrimination, privacy, cybersecurity, administrative burden, and labor displacement remain distinct outcomes that can interact with dependence. Biased outputs can be harder to challenge when review capacity is weak; privacy safeguards do not by themselves preserve task knowledge when AI performs core reasoning. Future studies can examine these interactions alongside public service quality and remedy.

8. Conclusion

Administrative judgment is at risk when public organizations substantially delegate core cognitive work while lacking the capacity required to review or resume it. The framework identifies this configuration as institutional AI dependence and assesses review and recovery separately. Repeated substitution can displace practice; authority and workflow defaults can reduce required challenge. Practice, corrective feedback, and knowledge retention shape later capacity, while continuing service obligations and limited alternatives can encourage further delegation. These mechanisms explain how initially useful delegation can become difficult to govern over time.

Public institutions keep delegation governable by maintaining qualified review, usable intervention powers, task–relevant practice, transferable knowledge, and tested recovery. The task’s legal duties and service consequences determine the required capacity. These arrangements make the ability to examine, contest, correct, and resume delegated work a continuing condition of AI-supported administration, providing a basis for current agentic governance and preparedness for more capable systems.

Declarations

Author Contributions

Chengwen Song: Conceptualization; Methodology; Investigation; Formal analysis; Visualization; Writing—original draft; Writing—review and editing; Project administration.

Funding

The author received no funding for this study.

Conflict of Interest

The author declares no conflicts of interest.

Acknowledgments

None.

Ethics Approval

This conceptual article analyses published literature and reports no newly collected human-participant data. Ethics committee approval is not applicable to this research design.

Informed Consent

Informed consent is not applicable because this study involved no newly recruited human participants.

Data Availability

This theoretical study draws on published literature cited in the reference list. No new empirical datasets were generated or analysed.

References

1. Alon-Barkat, S., & Busuioc, M. (2023). Human–AI interactions in public sector decision making: “Automation bias” and “selective adherence” to algorithmic advice. *Journal of Public Administration Research and Theory*, 33(1), 153–169. <https://doi.org/10.1093/jopart/muac007>
2. Anastasopoulos, L. J. (2026). Artificial intelligence, bureaucratic discretion, and democratic administration. *Public Administration Review*. Advance online publication. <https://doi.org/10.1111/puar.70201>
3. Argote, L., & Miron-Spektor, E. (2011). Organizational learning: From experience to knowledge. *Organization Science*, 22(5), 1123–1137. <https://doi.org/10.1287/orsc.1100.0621>
4. Bainbridge, L. (1983). Ironies of automation. *Automatica*, 19(6), 775–779. [https://doi.org/10.1016/0005-1098\(83\)90046-8](https://doi.org/10.1016/0005-1098(83)90046-8)
5. Benda, N. C., Novak, L. L., Reale, C., & Ancker, J. S. (2022). Trust in AI: Why we should be designing for appropriate reliance. *Journal of the American Medical Informatics Association*, 29(1), 207–212. <https://doi.org/10.1093/jamia/ocab238>
6. Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. *European Law Journal*, 13(4), 447–468. <https://doi.org/10.1111/j.1468-0386.2007.00378.x>
7. Bovens, M., & 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*, 62(2), 174–184. <https://doi.org/10.1111/0033-3352.00168>
8. Bullock, J. B., Young, M. M., & Wang, Y.-F. (2020). Artificial intelligence, bureaucratic form, and discretion in public service. *Information Polity*, 25(4), 491–506. <https://doi.org/10.3233/IP-200223>
9. Deng, Y., & Sun, Y. (2025). How does the usage of artificial intelligence affect felt administrative accountability of street-level bureaucrats? The mediating effect of perceived discretion. *Public Management Review*, 27(11), 2795–2815. <https://doi.org/10.1080/14719037.2024.2370982>

10. Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, 37(2), 381–394. <https://doi.org/10.1518/001872095779064555>
11. Green, B. (2022). The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review*, 45, 105681. <https://doi.org/10.1016/j.clsr.2022.105681>
12. Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, 10(1–2), 18–26. <https://doi.org/10.1007/s13162-020-00161-0>
13. Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1–22, Article 1121). Association for Computing Machinery. <https://doi.org/10.1145/3706598.3713778>
14. Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392
15. OECD. (2026a). *The agentic AI landscape and its conceptual foundations*. OECD Artificial Intelligence Papers, No. 56. OECD Publishing. <https://doi.org/10.1787/396cf758-en>
16. OECD. (2026b). *Government unstuck: The next frontier of government transformation*. OECD Publishing. <https://doi.org/10.1787/464c0cff-en>
17. Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230–253. <https://doi.org/10.1518/001872097778543886>
18. Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. <https://doi.org/10.1016/j.tics.2016.07.002>
19. Selten, F., Robeer, M., & Grimmelikhuijsen, S. (2023). “Just like I thought”: Street-level bureaucrats trust AI recommendations if they confirm their professional judgment. *Public Administration Review*, 83(2), 263–278. <https://doi.org/10.1111/puar.13602>
20. Song, C., Hu, Y., Chen, F., Lenz, S., Li, J., Feng, Y., Zeng, X., & Wu, M. (2026). When algorithms become sacred: A cognitive-psychological model of trust, authority, and meaning in AI interaction. *Journal of Social Cognition and Communication*, 1(1), 22–31. <https://doi.org/10.63802/jssc.V1.I1.291>
21. Wong, S.-M. (2026). Office and unowned architecture: Public judgment in the age of algorithmic governance. *Ethics and Information Technology*, 28, Article 48. <https://doi.org/10.1007/s10676-026-09919-9>
22. Wuttke, A., Jungherr, A., & Rauchfleisch, A. (2026). Artificial intelligence in government: Why people feel they lose control. *Journal of European Public Policy*. Advance online publication. <https://doi.org/10.1080/13501763.2026.2696304>
23. Young, M. M., Bullock, J. B., & Lecy, J. D. (2019). Artificial discretion as a tool of governance: A framework for understanding the impact of artificial intelligence on public administration. *Perspectives on Public Management and Governance*, 2(4), 301–313. <https://doi.org/10.1093/ppmgov/gvz014>
24. Zhang, Y., Hsieh, C.-W., & Liu, F. (2025). Employee response to GenAI: Exploring the effect of ChatGPT dependence on work effort in a public service context. *Public Performance & Management Review*, 48(6), 1302–1330. <https://doi.org/10.1080/15309576.2025.2537107>