

# When AI Policies Fail in Practice: Shadow AI as a Structural Policy– Practice Governance Misalignment

Mia Wilson\*<sup>1</sup>, Ethan Moore<sup>2</sup>

Email: [mia998765545@gmail.com](mailto:mia998765545@gmail.com)

<sup>1</sup>National University of Singapore, Singapore

\*Corresponding Author

## Abstract

*The widespread adoption of Artificial Intelligence (AI) has led organizations to establish formal governance frameworks aimed at mitigating ethical, legal, and operational risks. Despite these efforts, AI governance frequently fails in practice, as evidenced by the growing prevalence of Shadow AI, defined as the unsanctioned use of AI tools by employees. Existing scholarly and practitioner discourses predominantly frame this phenomenon as a compliance failure or security vulnerability, thereby emphasizing stricter controls and enhanced employee training as primary remedies. This conceptual study challenges that prevailing view by arguing that Shadow AI represents a structural manifestation of policy–practice misalignment rather than a problem of individual deviance. The study develops a diagnostic framework that identifies three constitutive dimensions of misalignment: temporal gaps (mismatches between governance processes and operational speed), utility gaps (misalignment between sanctioned tools and task-specific needs), and autonomy–control gaps (tensions between professional discretion and standardization). Drawing on a theory-driven conceptual methodology integrating sociotechnical systems theory with policy–practice analysis, and illustrated through structured synthetic organizational scenarios, the study demonstrates how governance designs that overlook the realities of situated work systematically generate Shadow AI practices. The analysis further suggests that adaptive governance models incorporating structured flexibility such as curated AI tool marketplaces and expedited approval pathways are theoretically more effective than highly rigid governance regimes. The primary contribution lies in advancing a practice-aware AI governance model that reframes Shadow AI as a diagnostic signal of systemic design flaws and provides a foundation for more legitimate and responsive AI governance.*

**Keywords:** *AI Governance; Adaptive Governance; Policy-Practice Gap; Shadow AI; Sociotechnical Systems.*

*Received in April 2026; Revised in May 2026; Accepted in July 2026; Published in August 2026*

## I. INTRODUCTION

The rapid integration of Artificial Intelligence (AI) into core organizational functions represents one of the most significant technological transformations of the digital era. From automating routine tasks to enabling sophisticated data analytics and decision support, AI promises substantial gains in efficiency, innovation, and competitive advantage (Birkstedt et al., 2023). In response to this transformative potential and its concomitant risks a complex ecosystem of AI governance has emerged. This ecosystem comprises ethical guidelines, risk management protocols, compliance standards, and technical safeguards designed to ensure responsible, secure, and aligned AI use (Attard-Frost et al., 2024; Schmitt, 2022). The governance imperative reflects a critical consensus: without effective oversight, AI deployment can lead to significant ethical breaches, legal liabilities, security vulnerabilities, and reputational damage (Daly et al., 2021). However, despite the maturity and proliferation of these governance mechanisms, their

effectiveness in real organizational contexts remains uneven, indicating a persistent gap between governance design and governance outcomes.

Despite the proliferation of formal AI policies, a persistent and troubling organizational reality challenges their efficacy: the widespread, covert use of unsanctioned AI tools by employees, a phenomenon termed "Shadow AI" (Puthal et al., 2025). Whether using a public large language model for drafting sensitive reports, an unvetted analytics tool for processing proprietary data, or an open-source library for developing features, employees are bypassing official channels at scale. The prevailing literature predominantly characterizes Shadow AI as a problem of individual non-compliance, a security blind spot, or an ethical lapse (Yu et al., 2024). This dominant narrative logically prescribes solutions centered on enhanced technological surveillance, stricter access controls, and more comprehensive compliance training. However, this perspective is critically limited. It treats the symptom as the disease, overlooking the possibility that Shadow AI is not an anomaly but a predictable, structural outcome a symptom of a deeper failure in governance system design. Moreover, existing studies have not sufficiently distinguished Shadow AI from related constructs such as Shadow IT, nor have they fully explained why similar governance approaches fail to prevent such practices across different organizational settings.

This study posits that the root cause of Shadow AI lies not in employee malfeasance but in a fundamental policy–practice misalignment. We argue that Shadow AI emerges systematically when the design logic of top-down, formal AI governance policies becomes structurally disconnected from the practical logic of bottom-up, everyday work (Martin et al., 2024; Waters-Lynch et al., 2025). Governance is often conceived as a static, design-time artifact (intended governance), but it is enacted, interpreted, and negotiated in the dynamic, messy context of practice (enacted governance). When the gap between the two becomes too wide, Shadow AI flourishes as a rational, albeit risky, employee adaptation strategy. In this sense, Shadow AI should be understood not merely as deviant behavior but as an adaptive response to misaligned institutional structures and operational constraints within organizations.

To investigate this core proposition, we address the following research question: How do structural misalignments between intended AI governance policies and enacted workplace practices systematically lead to the emergence of Shadow AI? The objectives of this conceptual study are threefold: (1) to reconceptualize Shadow AI from a compliance problem to a diagnostic indicator of systemic governance failure; (2) to develop an analytical framework that delineates the specific dimensions and mechanisms of this policy–practice misalignment; and (3) to propose a paradigm shift toward adaptive, practice-aware governance models. Theoretically, this research bridges and extends discrete scholarly domains. It integrates literature from AI governance

(focused on design), studies of policy implementation gaps, and sociotechnical systems theory into a novel, cohesive framework. By introducing a governance-in-practice lens, we shift analytical attention from the design of rules to their enactment, adaptation, and subversion in daily organizational life, offering a more nuanced explanation for technology governance failures, drawing on insights from sociotechnical transitions in other industries (Chung et al., 2023; Sovacool et al., 2021) and principles of polycentric governance (Keudel & Huss, 2024). Importantly, this study advances the literature by proposing an explanatory model rather than a purely descriptive conceptual discussion, thereby providing a structured account of the causal mechanisms through which governance misalignment produces Shadow AI. Furthermore, the proposed framework contributes a novel integration of AI governance and sociotechnical perspectives that extends beyond existing Shadow IT and compliance-oriented studies by explicitly modeling the interaction between formal policy structures and situated workplace practices.

Practically, this study provides organizational leaders, policymakers, and IT governance professionals with a pragmatic diagnostic toolkit. The framework enables them to identify latent friction points and weaknesses in their AI governance architectures. Furthermore, it offers evidence-based guidance for designing more resilient and legitimate governance that balances necessary control with the flexibility required for innovation and operational responsiveness. By doing so, the study contributes not only to improved governance design but also to a deeper understanding of why conventional enforcement-centric approaches are insufficient in addressing Shadow AI. In addition, it positions governance as an evolving, practice-sensitive process rather than a static compliance mechanism, encouraging organizations to adopt adaptive strategies that better align policy intent with operational realities. The remainder of this paper is structured as follows. Section II reviews relevant literature and synthesizes the identified research gap. Section III details our conceptual methodology. Section IV presents the core conceptual framework and illustrative findings. Section V discusses the theoretical and managerial implications. Section VI acknowledges limitations and suggests future research directions, and Section VII concludes.

## **II. LITERATURE REVIEW**

### *A. AI Governance: The Design-Logic Perspective*

Organizational AI governance is formally defined as "the integrated ensemble of structures, formal processes, and evaluative standards established to strategically direct and control the adoption and use of AI technologies to achieve organizational objectives while mitigating associated risks" (Mäntymäki et al., 2022). This burgeoning field is predominantly concerned with design: formulating ethical principles (e.g., fairness, accountability, transparency),

constructing risk taxonomies, and establishing control mechanisms like review boards, audit trails, and model validation (Batoool et al., 2025; Roberts et al., 2024). The underlying perspective is often top-down and techno-centric, prioritizing ex-ante planning, standardization, and risk mitigation. It embodies what we term the design logic a focus on creating an ideal, formal structure intended to produce orderly and compliant outcomes (Daly et al., 2021). While essential for establishing a baseline of responsibility, this perspective can inadvertently treat governance as a static blueprint, underestimating the complexity of its implementation in diverse organizational contexts, a challenge also recognized in technical AI governance literature (Reuel et al., 2025) and studies of policy failure mitigation (Leong & Howlett, 2022). Notably, although this design-oriented literature provides robust normative and procedural guidance, it largely abstracts away from how governance mechanisms operate once embedded in situated organizational practices, thereby limiting its explanatory power regarding real-world deviations.

### *B. The Practice-Logic of Technology Adoption and Shadow IT*

In stark contrast to the design-oriented governance literature, a long-standing body of research on technology adoption illuminates the practice logic that governs how tools are actually used. Studies of "Shadow IT" the unsanctioned acquisition and use of technology by business units demonstrate that employees will routinely bypass formal IT channels when those channels are perceived as slow, unresponsive, or offering inferior tools (Marpu et al., 2025). This behavior is often motivated by a genuine desire for efficiency, innovation, and personal productivity, not mere rebellion. This aligns with a practice-based view of technology, which shifts scholarly focus from the design of artifacts to how they are appropriated, adapted, and made meaningful within the specific contexts of everyday work. Emerging empirical work on Shadow AI diligently catalogs its associated risks data leakage, model poisoning, copyright infringement, and security breaches (Puthal et al., 2025; Yu et al., 2024). However, it largely replicates the "compliance failure" frame of earlier Shadow IT research, focusing on the manifestations of the problem rather than its systemic antecedents. As a result, existing studies explain what Shadow AI is and what risks it poses, but fall short of explaining why such behaviors persist even in organizations with formalized and seemingly comprehensive governance structures.

### *C. Bridging the Gap: Sociotechnical Systems and Policy-Practice Misalignment*

Sociotechnical systems (STS) theory provides the crucial integrative lens needed to bridge the design-practice divide. STS posits that optimal organizational performance and sustainability arise from the joint optimization of the social subsystem (people, skills, values, relationships) and the technical subsystem (tools, processes, tasks) (Andersson et al., 2021; Kim et al., 2022). A core STS tenet is that imposing a rigid technical system without regard for social dynamics leads to

workarounds, low morale, and suboptimal outcomes a pattern that perfectly describes the genesis of Shadow AI. This perspective is reinforced by systematic reviews of sociotechnical transitions in various industries, which highlight the recurrent tension between technological design and social practice in sectors like ceramics and glass manufacturing (Furszyfer Del Rio et al., 2022a, 2022b), and is further evidenced in studies of value inference within sociotechnical systems (Liscio et al., 2023). The concept of policy-practice gaps is well-established in public administration, management, and implementation science (Troje, 2021; Whiteside, 2021). This literature rigorously documents how formal policies are invariably transformed during implementation by local contexts, resource constraints, professional discretion, and interpretive agency. Studies in domains from sustainability to inter-organizational collaboration show that rigid, top-down policies often fail, generating unintended consequences and widespread workarounds (Ingstrup et al., 2021; Steinhall et al., 2023). Nevertheless, while STS and policy-implementation literatures acknowledge misalignment between design and practice, they do not explicitly formalize how such misalignment translates into specific behavioral patterns like Shadow AI within contemporary AI governance contexts.

#### *D. Synthesis and Identified Research Gap*

A critical synthesis reveals a pronounced gap at the intersection of these literatures within the AI governance domain. The existing body of work on AI governance provides detailed prescriptions on how to design control mechanisms, principles, and oversight structures, yet it remains largely silent on how these mechanisms behave when enacted in complex organizational environments. Parallel to this, the Shadow AI and Shadow IT literatures document the persistence of unsanctioned technology use but primarily interpret it as a compliance issue rather than as an outcome shaped by governance design itself. In addition, the policy-practice gap literature and sociotechnical systems theory offer strong conceptual foundations for understanding implementation dynamics and system interactions, yet these perspectives have not been systematically integrated into AI governance to explain emergent shadow behaviors. Crucially, no existing stream of literature sufficiently integrates these three domains into a unified, mechanism-based explanation that links governance design attributes, implementation contexts, and user adaptations in a causal chain. As a result, the current discourse treats Shadow AI as a behavioral output to be controlled and governance as a design input, without sufficiently theorizing their recursive, co-constitutive relationship. This fragmentation across literatures creates a theoretical blind spot: while each stream is individually well-developed, their lack of integration prevents a comprehensive understanding of how and why governance systems systematically generate unintended shadow practices. The missing element is therefore a coherent, explanatory framework that elucidates the specific mechanisms through which the

attributes of AI governance policies can structurally and predictably induce the shadow practices they aim to prevent. This study addresses this explanatory gap by developing a policy-practice misalignment framework for AI governance, thereby synthesizing insights across these previously disconnected literatures into a unified causal model.

### III. RESEARCH METHODOLOGY

#### *A. Research Design*

This study employs a conceptual research design aimed at theory elaboration and integrative framework development (Birkstedt et al., 2023; Mäntymäki et al., 2022). The primary goal is the construction of a logically coherent, theoretically grounded model to explain the systemic origins of Shadow AI, rather than conducting empirical hypothesis testing. This methodological approach is particularly appropriate for investigating complex and emerging organizational phenomena where well-established causal relationships have not yet been clearly defined within the literature. By structuring the conceptual field and generating a set of testable propositions, this study establishes a necessary foundation to guide and inform subsequent empirical research efforts, following methodological approaches used in desk-based analytical case studies (Nugroho & Wibowo, 2025). Accordingly, the design prioritizes theoretical integration and mechanism-building over empirical generalization, positioning the study as an explanatory contribution that formalizes relationships across previously fragmented literatures.

#### *B. Analytical Procedure: A Structured Four-Stage Process*

Our analytical procedure followed a rigorous, four-stage process, which is comprehensively visualized in Figure 1. The initial stage involved systematically extracting the core imperatives and assumptions of intended governance from the AI governance literature, identifying logics such as risk mitigation and ex-ante control. Subsequently, we distilled the competing practice logics from literature on Shadow IT and sociotechnical systems, focusing on imperatives like operational agility and professional autonomy. We then conducted a comparative contraposition of these two sets of logics, which revealed fundamental tensions and allowed us to inductively derive the core dimensions of misalignment and develop the AI Governance Adaptation Cycle. Finally, to elaborate and explore the framework's implications, we constructed synthetic organizational scenarios as heuristic tools to trace logical consequence pathways and theorize about potential governance parameter spaces. Across these stages, the analysis followed a structured process of conceptual coding, thematic synthesis, iterative abstraction, and comparative mapping to ensure that extracted logics were systematically categorized, contrasted, and synthesized into higher-order analytical constructs.

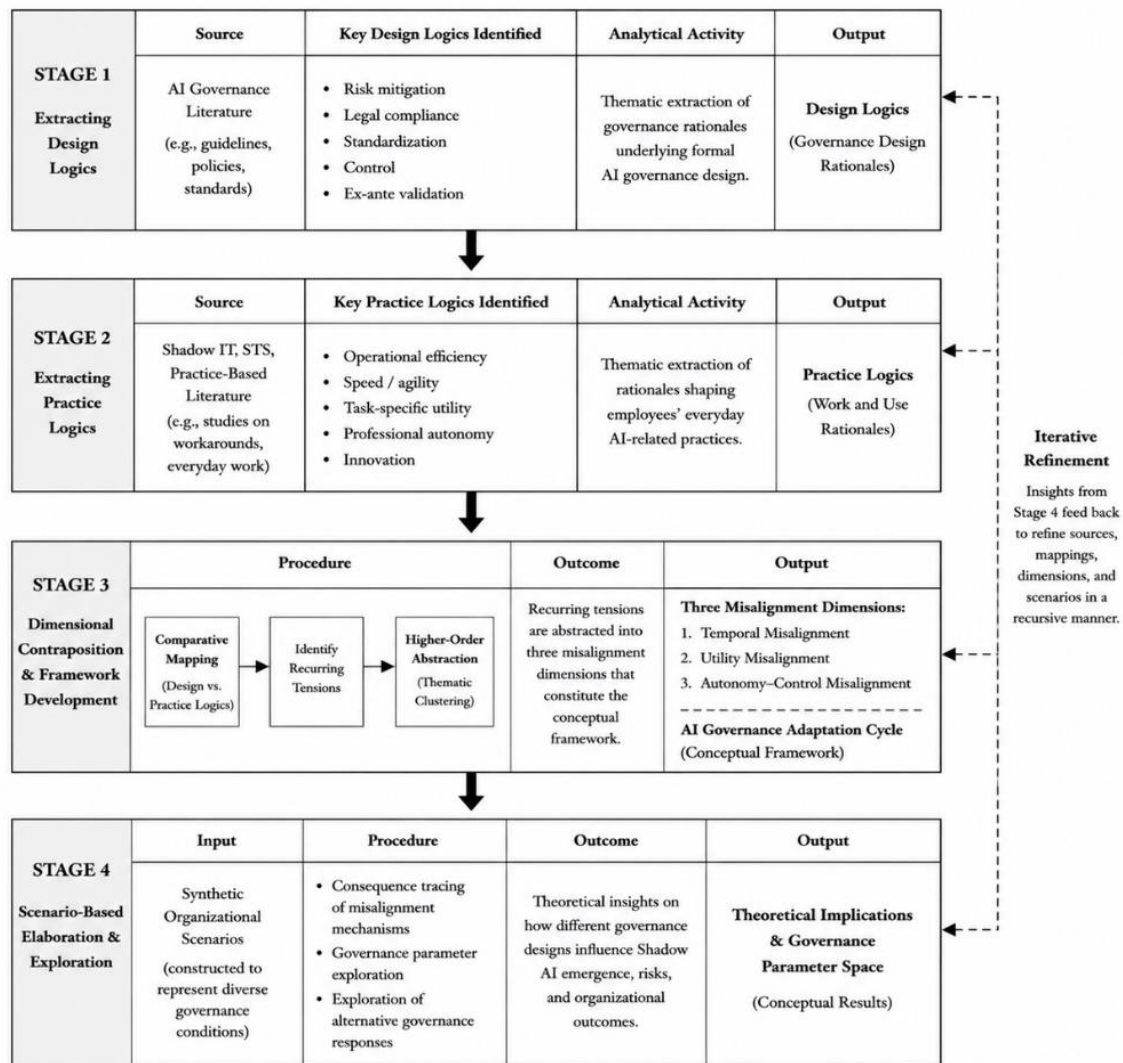
Stage 1: Extracting Design Logics. We systematically analyzed the literature on AI governance and its formal, design-oriented principles to identify the core imperatives, assumptions, and formal structures of intended governance. Key extracted logics included: risk mitigation, legal compliance, standardization, control, and ex-ante validation (Daly et al., 2021; Mäntymäki et al., 2022). This foundational step established the formal, top-down logic that characterizes much of the prevailing governance discourse and its focus on creating an ideal, controlled structure for AI use. The extraction process involved iterative reading and conceptual coding of relevant texts, followed by clustering similar governance principles into coherent thematic categories representing dominant design logics.

Stage 2: Extracting Practice Logics. We distilled the competing logics from literature on Shadow IT, the practice-based view of technology, and sociotechnical systems theory, which focus on how tools are actually adopted and used in daily work. Key extracted logics included: operational efficiency, speed/agility, task-specific utility, professional autonomy, and innovation (Andersson et al., 2021; Marpu et al., 2025; Waters-Lynch et al., 2025). This stage captured the bottom-up, situated rationalities that drive technology adoption and adaptation in everyday work contexts, often in contrast to formal policy mandates. Similar to Stage 1, we applied iterative abstraction and thematic grouping to organize practice-oriented behaviors and motivations into a structured set of recurring logics that reflect user-driven adaptation patterns.

Stage 3: Dimensional Contraposition and Framework Development. Through comparative analysis, we juxtaposed the design logics from Stage 1 against the practice logics from Stage 2. This systematic contraposition revealed fundamental, recurring tensions between the imperatives of control and the needs of practice. By synthesizing these tensions, we inductively derived the core dimensions of misalignment where policy and practice consistently clash. Specifically, we employed a comparative mapping procedure in which each design logic was aligned against corresponding practice logics to identify points of conflict, overlap, or incompatibility, followed by clustering these tensions into higher-order categories that constitute the derived dimensions of misalignment. This critical stage produced the core conceptual framework: the dimensions of misalignment and the dynamic AI Governance Adaptation Cycle, which models the relationship between governance design, practice, and shadow outcomes.

Stage 4: Scenario-Based Elaboration and Exploration. To illustrate the framework's operation, test its internal logic, and explore its implications, we constructed synthetic organizational scenarios. These are heuristic, analytical constructs, not empirical data, and they represent plausible, recognizable organizational contexts (e.g., a marketing department with rapid campaign cycles versus a legal department with careful review processes). We used these scenarios to "run

simulations" of the framework, tracing how different configurations of policy rigidity and practice needs would logically lead to varying propensities for Shadow AI. It is important to emphasize that these scenarios are not intended to represent empirical observations or support statistical generalization, but rather serve as controlled thought experiments designed to elucidate causal mechanisms and evaluate the internal consistency of the proposed framework. This process allowed us to explore consequence pathways, test the framework's robustness against edge cases, and identify potential parameter spaces for effective, adaptive governance design, as summarized in the findings section.



**Note:** The process proceeds iteratively. Insights from Stage 4 (theoretical implications and parameter space) feed back to refine the comparative mappings, dimensions, and scenarios used in earlier stages.

**Figure 1. Conceptual Research Methodology Process**

Conceptual research methodology illustrating a linear four-stage analytical process, from the extraction of governance design logics and practice logics to framework development and scenario-based theoretical exploration.

### *C. Validity and Robustness*

The validity and robustness of our conceptual framework are supported through multiple deliberate considerations. Construct validity is firmly anchored in the framework's explicit grounding within established mid-range theories, including sociotechnical systems theory and policy implementation studies (Kim et al., 2022; Troje, 2021). We ensured internal validity, or logical consistency, through iterative reasoning, a structured peer debriefing process focused on the contraposition stage, and systematic testing of the framework against edge cases within our synthetic scenarios. While traditional empirical reliability metrics are not directly applicable, we pursued analytical transparency and rigor by meticulously documenting each step of our procedural logic.

The credibility of the conceptual approach is further reinforced through cross-literature triangulation, whereby insights from AI governance, Shadow IT, and sociotechnical systems literatures converge to support the same underlying mechanisms of policy–practice misalignment. This documentation allows for the scrutiny and potential replication of our analytical process by other scholars, thereby upholding the standards of rigorous conceptual research. In addition, the iterative and theory-grounded nature of the analysis ensures that the derived framework is not arbitrary, but instead emerges from systematically synthesized and theoretically consistent interpretations across multiple established bodies of literature.

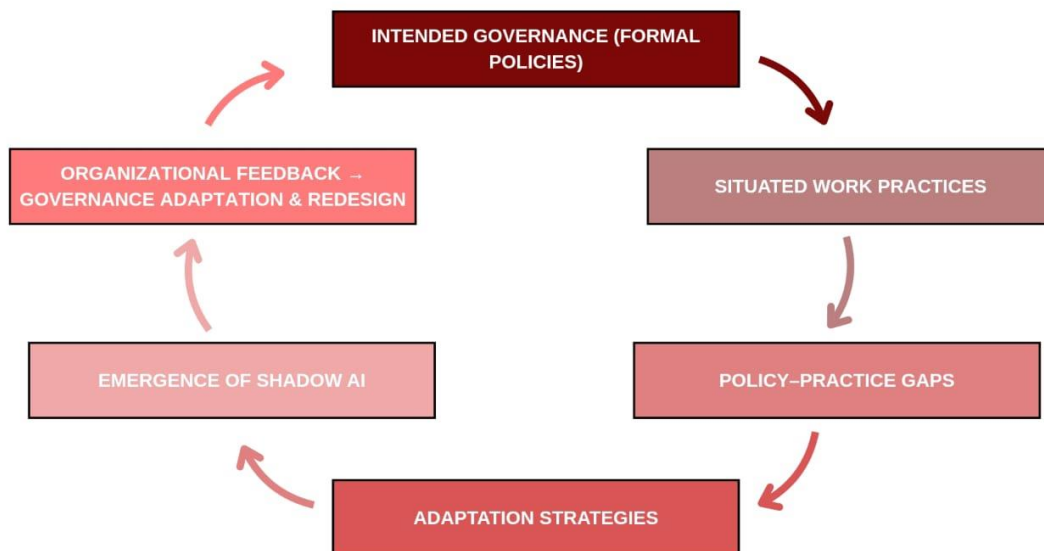
### *D. Ethical Considerations*

As a purely conceptual and non-empirical study, the ethical considerations of this research primarily concern standards of scholarly integrity, analytical rigor, and transparent theoretical reasoning. The study ensures the accurate representation, critical interpretation, and proper attribution of all prior scholarly work that informs the development of the conceptual framework. Scenario-based illustrations are employed solely as analytical heuristics to clarify theoretical mechanisms and relationships, rather than to produce empirical claims or normative prescriptions. By clearly distinguishing conceptual reasoning from empirical inference, the study seeks to prevent misinterpretation and uphold ethical standards in the communication of theoretical research. Furthermore, all analytical steps are transparently documented to ensure intellectual honesty, allowing readers to trace how conclusions are derived from existing literature rather than from unstated assumptions or empirical generalizations.

#### IV. RESULT/FINDINGS AND DISCUSSION

##### A. Result

All results, tables, and figures presented in this section are derived from conceptual reasoning and scenario-based analytical exploration. They are not based on empirical measurement, datasets, or statistical inference. The scenarios, numerical values, and organizational illustrations are heuristic constructs used solely to demonstrate the internal logic and implications of the proposed framework. Accordingly, all findings should be interpreted as theoretically grounded illustrations intended to explain mechanisms and relationships, rather than as empirically validated outcomes or generalizable quantitative results.



**Figure 2. The AI Governance Adaptation Cycle**

##### a) Core Conceptual Model: The AI Governance Adaptation Cycle

The central output of this research is a dynamic model reframing the relationship between governance, practice, and Shadow AI (Figure 2). The model depicts a continuous cycle where formal Intended Governance interacts dynamically with Situated Work Practices, revealing specific Policy–Practice Gaps that prompt various Employee Adaptation Strategies. A primary, though risky, adaptation is the Emergence of Shadow AI, which then generates Organizational Feedback in the form of incidents, inefficiencies, or unexpected innovations. This feedback critically informs Governance Adaptation & Redesign, leading to revised policies that restart the cycle anew. This cyclical model repositions Shadow AI not as a terminal compliance failure but

as a critical feedback mechanism within the organizational learning system; failure to learn from and respond to this feedback perpetuates and deepens structural misalignment.

Importantly, this model should be interpreted as a conceptual representation of systemic dynamics rather than a measurable empirical process, emphasizing causal relationships inferred from literature synthesis and theoretical integration. Figure 2 illustrates the AI governance adaptation cycle, showing how intended AI governance (formal policies) interacts with situated work practices. Misalignments between policy design and everyday work generate policy–practice gaps that prompt local adaptation strategies, including the emergence of Shadow AI. Rather than representing a mere compliance failure, Shadow AI is conceptualized as an organizational feedback mechanism that reveals structural weaknesses in governance design and informs subsequent governance adaptation and redesign.

**Table 1. Dimensions of AI Governance Policy–Practice Misalignment**

| Dimension                     | Definition                                                                                                                                                        | Key Indicators                                                                                                                  | Primary Organizational Impact                                                                                                                  |
|-------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|
| Temporal Misalignment         | A disconnect between the speed of governance processes and the operational tempo of work.                                                                         | Long approval cycles, slow tool updates, infrequent policy reviews vs. rapid project sprints, fast-moving markets.              | Creates pressure for shortcuts; most acute in units like Marketing, Software Development.                                                      |
| Utility Misalignment          | A gap between the functionality, ease of use, or relevance of sanctioned AI tools and available unsanctioned alternatives.                                        | Sanctioned tools are outdated, lack key features, have poor UX, or are overly generic for specialized tasks.                    | Drives tool choice based on efficacy over compliance; critical in R&D, Data Science.                                                           |
| Autonomy–Control Misalignment | A tension between employees' need for professional discretion in performing their work and governance requirements for standardization and centralized oversight. | Highly restrictive whitelists, inability to customize workflows, lack of input in tool selection, "one-size-fits-all" policies. | Encourages resistance to perceived deskilling; particularly salient among knowledge professionals (e.g., lawyers, researchers, and engineers). |

*b) The Three Dimensions of Misalignment*

The systematic contraposition of design and practice logics crystallized into three constitutive dimensions of policy-practice misalignment, each representing a distinct channel through which governance failure predictably manifests (Table 1). Temporal Misalignment captures the critical disconnect between the slow speed of governance processes and the rapid operational tempo of modern work, creating immense pressure for shortcuts in time-sensitive units. Utility Misalignment addresses the persistent gap between the functionality of sanctioned AI tools and available unsanctioned alternatives, driving tool selection based on efficacy rather than compliance in specialized domains. Autonomy-Control Misalignment encapsulates the

fundamental tension between employees' need for professional discretion and governance's push for standardization, provoking resistance among knowledge workers who perceive over-control as deskilling. These three dimensions are analytically distinct but often coexist and reinforce each other in practice, creating a powerful, cumulative impetus for the proliferation of Shadow AI across organizational units. These dimensions are not empirically measured variables but theoretically derived constructs obtained through iterative abstraction and comparative synthesis of recurring tensions across the analyzed literatures.

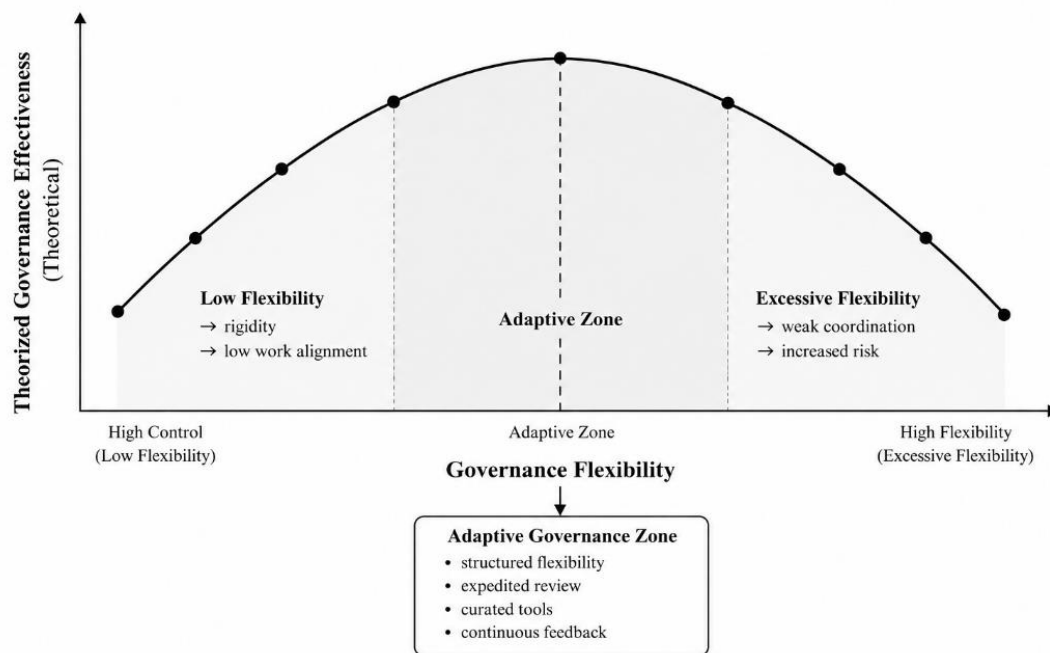
*c) Illustrative Insights from Scenario Exploration*

Applying the diagnostic framework to synthetic organizational scenarios yields concrete, illustrative insights into its explanatory power and practical utility. First, scenario analysis reveals significant variation in Shadow AI propensity across different organizational units, directly correlated with their primary type of misalignment (Table 2). For instance, the Marketing department's extreme temporal misalignment with a 60-day policy process for 3-day tasks logically explains its high theoretical propensity for shadow tools, while Legal's more aligned timelines correspond to lower temporal pressure but heightened autonomy-control tensions. Second, exploring the parameter space of governance design suggests a non-linear, inverted U-shaped relationship between control and overall effectiveness, as visualized in Figure 3.

This analysis identifies a theorized "Adaptive Governance Zone" where moderate, structured flexibility optimizes the trade-off between alignment, risk, and innovation, outperforming both highly rigid and excessively lax approaches across key metrics (Table 3). These illustrative findings demonstrate how the framework can diagnose root causes of Shadow AI and guide the design of more effective, responsive governance regimes. Again, these scenario outputs are interpretive and illustrative, intended to demonstrate logical consistency and mechanism behavior rather than to represent real-world empirical distributions or predictive benchmarks.

**Table 2. Illustrative Shadow AI Propensity Across Organizational Units (Synthetic Scenario Output)**

| Organizational Unit    | Policy Process Duration | Work Cycle Requirement | Primary Misalignment Dimension | Illustrative Shadow AI Propensity* |
|------------------------|-------------------------|------------------------|--------------------------------|------------------------------------|
| Marketing              | 60 days                 | 3 days                 | Temporal                       | 92%                                |
| Software Development   | 45 days                 | 7 days                 | Temporal, Utility              | 78%                                |
| Research & Development | 90 days                 | 30 days                | Utility, Autonomy-Control      | 65%                                |
| Legal & Compliance     | 21 days                 | 21 days                | Autonomy-Control               | 35%                                |



**Note:** The inverted-U relationship represents a theoretical proposition derived from scenario-based conceptual analysis. Governance effectiveness is theorized to be highest within an adaptive governance zone. The curve is intended as a conceptual illustration and should not be interpreted as an empirical estimate.

**Figure 3. Theorized Governance Flexibility–Effectiveness Relationship**

This figure presents a theorized relationship between the degree of AI governance flexibility and organizational performance. Highly rigid governance constrains work practices and innovation, resulting in suboptimal performance. Conversely, excessively flexible or weakly structured governance increases risk exposure and coordination failures, also reducing performance. Optimal performance is achieved under adaptive governance, where sufficient flexibility allows alignment with situated work practices while maintaining necessary oversight and control. The figure conceptualizes governance effectiveness as contingent on balance rather than maximal control or unrestricted flexibility.

**Table 3. Illustrative Comparison of Governance Design Models (Synthetic Scenario Output)**

| Design Parameter                       | Minimal Flexibility (Strict Control) | Moderate Flexibility (Adaptive Zone)   | Excessive Flexibility (Lax Control) |
|----------------------------------------|--------------------------------------|----------------------------------------|-------------------------------------|
| Approval Model                         | Central committee, full audit        | Curated marketplace + expedited review | No approval; post-hoc audit only    |
| Typical Timeframe                      | 30-90 days                           | 7-14 days                              | <24 hours                           |
| Illustrative Propensity for Shadow AI* | High (85-95%)                        | Low-Moderate (25-35%)                  | Moderate-High (60-70%)              |
| Theorized Alignment Score (1-5)*       | Low (2.1)                            | High (4.2)                             | Moderate (2.8)                      |
| Theorized Risk Score (1-5)*            | Low (1.8)                            | Moderate (2.5)                         | High (4.5)                          |

## B. Discussion

*a) Shadow AI as a Systemic Feedback Mechanism*

The findings of this conceptual study reposition Shadow AI from a peripheral compliance anomaly to a central diagnostic signal of structural weaknesses in AI governance design. The AI Governance Adaptation Cycle (Figure 2) demonstrates that Shadow AI emerges not randomly, but predictably, when formal governance architectures fail to align with the temporal, functional, and professional realities of situated work practices. This finding extends prior studies that frame Shadow AI primarily as a security exposure or ethical risk (Puthal et al., 2025; Yu et al., 2024) by showing that such risks are downstream manifestations of governance misalignment rather than isolated employee misconduct.

The three dimensions of misalignment identified in Table 1 provide analytical specificity to this argument. Temporal misalignment explains why Shadow AI is disproportionately concentrated in units characterized by rapid execution cycles, such as marketing and software development, where lengthy approval processes fundamentally contradict operational demands. This observation is consistent with earlier Shadow IT research demonstrating that speed asymmetries between formal controls and work rhythms systematically generate workarounds (Marpu et al., 2025), while extending this explanation to AI governance contexts in which decision latency may undermine organizational responsiveness. Importantly, the model suggests that suppressing Shadow AI through stricter enforcement does not eliminate the underlying misalignment; instead, it may obscure feedback signals that could inform governance adaptation. Accordingly, Shadow AI should be interpreted as an endogenous indicator of governance deficiencies that require structural recalibration rather than solely behavioral correction.

*b) Advancing a Practice-Aware Perspective in AI Governance Theory*

From a theoretical standpoint, this study advances AI governance scholarship by introducing a policy–practice misalignment lens that explains how governance failures emerge in practice rather than merely prescribing ideal governance principles. While existing AI governance literature frequently emphasizes normative frameworks centered on ethics, risk management, and accountability (Batoool et al., 2025; Mäntymäki et al., 2022), it often assumes that compliance will follow once appropriate rules and tools are established.

The present findings challenge this assumption by demonstrating that governance effectiveness depends on its alignment with everyday work practices. The utility and autonomy–control dimensions are particularly important in extending sociotechnical systems (STS) theory into AI governance. Consistent with STS scholarship, which emphasizes the joint optimization of social and technical subsystems (Andersson et al., 2021), the findings indicate that governance regimes prioritizing technical control at the expense of professional discretion and task-specific

functionality may unintentionally encourage shadow practices. This insight also aligns with policy implementation research showing that rigid, top-down policies are frequently reshaped or bypassed during enactment (Ingstrup et al., 2021; Troje, 2021). Rather than treating these perspectives separately, the present study integrates them into a unified explanatory mechanism that links temporal, utility, and autonomy–control misalignments to the emergence of Shadow AI.

*c) Theoretical Contributions*

This study contributes to AI governance theory in three ways. First, it shifts the analytical focus from static rule design to the dynamic alignment between governance structures and workplace practices. Second, it extends policy–practice gap literature by operationalizing misalignment as a multidimensional construct rather than a generic implementation challenge. Third, within the sociotechnical systems perspective, it identifies specific pathways through which governance tensions generate adaptive behaviors such as Shadow AI. Taken together, these contributions provide a mechanism-based explanation of how governance design characteristics influence behavioral outcomes through identifiable forms of misalignment.

*d) Interpreting Scenario-Based Findings in Relation to Prior Research*

The scenario-based exploration further develops the analysis by illustrating how different configurations of governance rigidity generate varying propensities for Shadow AI across organizational contexts (Tables 2 and 3). Although illustrative rather than empirical, these scenarios align with patterns reported in studies of Shadow IT and digital workarounds, where moderate flexibility often performs better than either strict control or laissez-faire approaches (Waters-Lynch et al., 2025).

Similarly, the theorized inverted U-shaped relationship between governance flexibility and effectiveness (Figure 3) is consistent with findings from adaptive regulation and polycentric governance research (Keudel & Huss, 2024; Sovacool et al., 2021). The concept of an Adaptive Governance Zone synthesizes these insights into a coherent theoretical construct. Rather than advocating maximal control or unrestricted autonomy, the findings suggest that governance effectiveness is most likely to emerge from calibrated flexibility that accommodates contextual variation across organizational settings. From this perspective, Shadow AI functions as an indicator that governance arrangements have moved outside the adaptive range required to balance control and flexibility. The contribution therefore lies not in empirically validating these relationships, but in providing a structured conceptual explanation for why such patterns may emerge under conditions of governance misalignment.

*e) Managerial Implications: From Enforcement to Governance Redesign*

The discussion also yields managerial implications that extend beyond conventional compliance-oriented responses to Shadow AI. The findings suggest that organizations should treat instances of Shadow AI as diagnostic inputs rather than solely as violations requiring sanction. By systematically mapping Shadow AI occurrences onto the three dimensions of misalignment, decision-makers can determine whether the underlying problem arises from procedural latency, inadequate tool utility, or excessive standardization that restricts professional discretion. This diagnostic orientation supports targeted governance redesign rather than blanket restrictions. As illustrated in Table 3, governance models incorporating structured flexibility such as curated AI tool marketplaces, provisional approvals, and participatory oversight mechanisms are theoretically better positioned to reduce shadow practices while maintaining acceptable levels of risk.

Operationally, organizations can implement adaptive governance through mechanisms such as tiered approval workflows (e.g., rapid-track versus standard-track requests), centralized AI tool marketplaces with vetted and categorized tools, and continuous feedback loops through which user-reported workarounds are documented and used to refine governance policies. These recommendations are consistent with emerging research on adaptive and participatory governance, which highlights legitimacy, responsiveness, and organizational learning as key determinants of regulatory effectiveness (Chung et al., 2023; Keudel & Huss, 2024). In practical terms, organizations may establish governance dashboards that monitor approval-cycle duration, tool-adoption patterns, and workaround frequencies, thereby enabling continuous adjustment of governance parameters based on observed behavior rather than static assumptions. Through this shift from enforcement toward governance redesign, organizations can transform the drivers of Shadow AI into opportunities for sanctioned innovation and organizational learning.

## V. CONCLUSION AND RECOMMENDATION

This study concludes that Shadow AI is best understood as a systemic manifestation of structural misalignment between formal AI governance arrangements and situated work practices, rather than as a consequence of individual non-compliance. By developing a policy–practice misalignment framework, this research identifies three core dimensions temporal, utility, and autonomy–control misalignment that consistently generate shadow practices across diverse organizational contexts. The proposed AI Governance Adaptation Cycle further conceptualizes Shadow AI as a recurring feedback signal that reveals deficiencies in governance design and implementation. In terms of contribution, this study advances the literature by providing a structured conceptual model that integrates governance design logics with practice-based realities, thereby offering a more holistic explanation of why Shadow AI emerges and persists in

organizations. Through this reframing, the study shifts the analytical focus from suppressing unauthorized AI use toward understanding the underlying governance conditions that enable its emergence.

Based on these findings, the study recommends that organizations move away from purely enforcement-driven AI governance toward adaptive, practice-aware governance design. This requires systematically diagnosing misalignment patterns across organizational units and tailoring governance mechanisms to the specific temporal demands, functional requirements, and professional autonomy of different work contexts. By treating Shadow AI as an informative input rather than solely as a compliance violation, organizations can leverage shadow practices to improve governance responsiveness and legitimacy. Future research should empirically validate the proposed dimensions of misalignment, examine the adaptability of the AI Governance Adaptation Cycle across different industries, and explore how contextual variables influence the intensity and form of Shadow AI emergence. Ultimately, such an adaptive approach enables the development of AI governance regimes that are not only responsible and risk-aware, but also agile, context-sensitive, and more effective in supporting sustainable AI-enabled innovation.

#### AI Use Declaration

The authors declare that artificial intelligence tools, including ChatGPT, were used exclusively for language-related refinement to improve clarity, grammar, and coherence of the manuscript. All scholarly work, including research design, data collection, analysis, interpretation, and conclusion development, was carried out independently by the authors. The authors assume full responsibility for the originality and integrity of the study.

#### REFERENCES

- Andersson, J., Hellsmark, H., & Sandén, B. (2021). The outcomes of directionality: Towards a morphology of sociotechnical systems. *Environmental Innovation and Societal Transitions*, 40, 108–131. <https://doi.org/10.1016/j.eist.2021.06.008>
- Attard-Frost, B., Brandusescu, A., & Lyons, K. (2024). The governance of artificial intelligence in Canada: Findings and opportunities from a review of 84 AI governance initiatives. *Government Information Quarterly*, 41(2). <https://doi.org/10.1016/j.giq.2024.101929>
- Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: a systematic literature review. *AI and Ethics*, 5(3), 3265–3279. <https://doi.org/10.1007/s43681-024-00653-w>
- Birkstedt, T., Minkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: themes, knowledge gaps and future agendas. In *Internet Research* (Vol. 33, Issue 7, pp. 133–167). Emerald Publishing. <https://doi.org/10.1108/INTR-01-2022-0042>
- Chung, C., Kim, J., Sovacool, B. K., Griffiths, S., Bazilian, M., & Yang, M. (2023). Decarbonizing the chemical industry: A systematic review of sociotechnical systems,

- technological innovations, and policy options. In *Energy Research and Social Science* (Vol. 96). Elsevier Ltd. <https://doi.org/10.1016/j.erss.2023.102955>
- Daly, A., Hagendorff, T., Hui, L., Mann, M., Marda, V., Wagner, B., & Wang, W. W. (2021). AI, Governance and Ethics: Global Perspectives. In *Constitutional Challenges in the Algorithmic Society* (pp. 182–201). Cambridge University Press. <https://doi.org/10.1017/9781108914857.010>
- Furszyfer Del Rio, D. D., Sovacool, B. K., Foley, A. M., Griffiths, S., Bazilian, M., Kim, J., & Rooney, D. (2022a). Decarbonizing the ceramics industry: A systematic and critical review of policy options, developments and sociotechnical systems. In *Renewable and Sustainable Energy Reviews* (Vol. 157). Elsevier Ltd. <https://doi.org/10.1016/j.rser.2022.112081>
- Furszyfer Del Rio, D. D., Sovacool, B. K., Foley, A. M., Griffiths, S., Bazilian, M., Kim, J., & Rooney, D. (2022b). Decarbonizing the glass industry: A critical and systematic review of developments, sociotechnical systems and policy options. In *Renewable and Sustainable Energy Reviews* (Vol. 155). Elsevier Ltd. <https://doi.org/10.1016/j.rser.2021.111885>
- Ingstrup, M. B., Aarikka-Stenroos, L., & Adlin, N. (2021). When institutional logics meet: Alignment and misalignment in collaboration between academia and practitioners. *Industrial Marketing Management*, 92, 267–276. <https://doi.org/10.1016/j.indmarman.2020.01.004>
- Keudel, O., & Huss, O. (2024). Polycentric governance in practice: the case of Ukraine's decentralised crisis response during the Russo-Ukrainian war. *Journal of Public Finance and Public Choice*, 39(1 Special Issue), 10–35. <https://doi.org/10.1332/25156918Y2023D000000002>
- Kim, J., Sovacool, B. K., Bazilian, M., Griffiths, S., Lee, J., Yang, M., & Lee, J. (2022). Decarbonizing the iron and steel industry: A systematic review of sociotechnical systems, technological innovations, and policy options. In *Energy Research and Social Science* (Vol. 89). Elsevier Ltd. <https://doi.org/10.1016/j.erss.2022.102565>
- Leong, C., & Howlett, M. (2022). Policy Learning, Policy Failure, and the Mitigation of Policy Risks: Re-Thinking the Lessons of Policy Success and Failure. *Administration and Society*, 54(7), 1379–1401. <https://doi.org/10.1177/00953997211065344>
- Liscio, E., Lera-Leri, R., Bistaffa, F., Dobbe, R. I., Jonker, C. M., Lopez-Sanchez, M., Rodriguez-Aguilar, J. A., & Murukannaiah, P. K. (2023). *Value Inference in Sociotechnical Systems: Blue Sky Ideas Track*. [www.ifaamas.org](http://www.ifaamas.org)
- Mäntymäki, M., Minkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics*, 2(4), 603–609. <https://doi.org/10.1007/s43681-022-00143-x>
- Marpu, R. P., McNamara, K. J., & Gupta, P. (2025). *The AI Shadow War: SaaS vs. Edge Computing Architectures*. <http://arxiv.org/abs/2507.11545>

- Martin, J., Ellström, P. E., Wallo, A., & Elg, M. (2024). Bridging the policy-practice gap: a dual challenge of organizational learning. *Learning Organization*, 32(7), 18–34. <https://doi.org/10.1108/TLO-05-2023-0079>
- Nugroho, S. A. A., & Wibowo, A. (2025). Evaluating Digital Transformation within Integration Limitations using Desk-Based Analytical Case Study. *Journal of Technology Informatics and Engineering*, 4(2), 289–299. <https://doi.org/10.51903/jtie.v4i2.365>
- Puthal, D., Mishra, A. K., Mohanty, S. P., Longo, A., & Yeun, C. Y. (2025). Shadow AI: Cyber Security Implications, Opportunities and Challenges in the Unseen Frontier. *SN Computer Science*, 6(5). <https://doi.org/10.1007/s42979-025-03962-x>
- Reuel, A., Bucknall, B., Casper, S., Fist, T., Soder, L., Aarne, O., Hammond, L., Ibrahim, L., Chan, A., Wills, P., Anderljung, M., Garfinkel, B., Heim, L., Trask, A., Mukobi, G., Schaeffer, R., Baker, M., Hooker, S., Solaiman, I., ... Trager, R. (2025). *Open Problems in Technical AI Governance*. <http://arxiv.org/abs/2407.14981>
- Roberts, H., Hine, E., Taddeo, M., & Floridi, L. (2024). Global AI governance: barriers and pathways forward. *International Affairs*, 100(3), 1275–1286. <https://doi.org/10.1093/ia/iaae073>
- Schmitt, L. (2022). Mapping global AI governance: a nascent regime in a fragmented landscape. *AI and Ethics*, 2(2), 303–314. <https://doi.org/10.1007/s43681-021-00083-y>
- Sovacool, B. K., Bazilian, M., Griffiths, S., Kim, J., Foley, A., & Rooney, D. (2021). Decarbonizing the food and beverages industry: A critical and systematic review of developments, sociotechnical systems and policy options. In *Renewable and Sustainable Energy Reviews* (Vol. 143). Elsevier Ltd. <https://doi.org/10.1016/j.rser.2021.110856>
- Steinhall, N., Mcpettit, R., Bond, J., Parks, M., Khan, M., Sharfarz, D., Cabrera, L., & Cabrera, D. (2023). *June 2023 Wicked Solutions for Wicked Problems Wicked Solutions for Wicked Problems: Misalignment in Public Policy Wicked Solutions for Wicked Problems*.
- Troje, D. (2021). Policy in practice: social procurement policies in the swedish construction sector. *Sustainability (Switzerland)*, 13(14). <https://doi.org/10.3390/su13147621>
- Waters-Lynch, J., Allen, D. W. E., Potts, J., & Berg, C. (2025). Shadow user innovation: governing covert generative-AI use for dynamic-capability renewal. *Innovation: Organization and Management*. <https://doi.org/10.1080/14479338.2025.2519546>
- Whiteside, H. L. (2021). *Misalignment Between Teaching And Learning*.
- Yu, J., Yu, Y., Wang, X., Lin, Y., Yang, M., Qiao, Y., & Wang, F.-Y. (2024). *The Shadow of Fraud: The Emerging Danger of AI-powered Social Engineering and its Possible Cure*. <http://arxiv.org/abs/2407.15912>