SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

The Procurement-To-Architecture Translation Gap In Public Sector AI

Mihir Shah

Independent Researcher, USA

Abstract
Public sector artificial intelligence engagements fail in a structural way that current literature has not addressed directly. Procurement language designed to produce competitive bidding and contractor accountability does not translate cleanly into the architectural decisions that determine whether an implementation succeeds. Vendors win contracts on the basis of procurement-shaped proposals and then encounter architectural realities the contract language did not anticipate. The gap between the contract and the architecture is where many deployments appear to stall. This paper analyzes that gap from a vendor-side solution architecture perspective, drawing on practitioner experience translating public-sector procurement vehicles into deliverable technical scope. It identifies three recurring patterns of mistranslation, examines the structural causes that reproduce them, and proposes a reduction framework of three practices applicable at different stages of the procurement lifecycle. Sector-specific considerations for federal, state, and local vehicles are discussed, along with limitations and broader implications for AI governance. The argument suggests that procurement language shapes implementation outcomes in ways that warrant structural attention rather than ad hoc adjustment, and that modest changes on both buyer and seller sides may reduce avoidable failure.

Keywords: Public Sector AI; AI Procurement; Solution Architecture; AI Governance; Cooperative Purchasing; Contract-To-Implementation Gap.

1. Introduction

Public sector procurement systems were built to serve goals that predate modern artificial intelligence. Fair competition, accountability, and defensible spending sit at the center of how governments buy technology [1]. These goals remain legitimate, and no argument here disputes them. The procurement vehicles that encode them, however, were designed for categories of technology that behaved predictably once specified. A server, a license, or a maintenance contract could be described in advance and delivered against that description with little surprise. Digital transformation in government has since shifted the underlying technology faster than the surrounding contractual machinery has adapted [2]. Artificial intelligence sharpens the mismatch because its performance depends on conditions that a solicitation rarely fixes in advance, including data quality and integration state. The result is a translation problem sitting between what a contract says and what an architecture must actually do.

Guidance on governing agency use of AI has grown quickly [3], and a dedicated standard for procuring AI and automated decision systems now exists [4]. Both signals suggest that buyers increasingly recognize AI as a distinct procurement category rather than another software line item. Recognition, however, has not yet closed the operational gap. Oversight reviews find that agencies have begun implementing AI requirements while several key steps remain incomplete [5]. Policy can declare an intention long before the contracting apparatus can express that intention in enforceable terms. The distance between policy intent and delivered systems is precisely where vendors and architects actually work, and it is wider than the guidance implies. That

distance, and the specific mechanisms that widen it, is the subject of this paper. The framing is diagnostic before it is prescriptive.

The mismatch tends to appear in three recurring forms, each described more fully in the sections that follow. Requirements often specify outcomes without specifying the architectural conditions on which those outcomes depend. Evaluation frequently rewards proposal quality rather than architectural feasibility against the buyer's real environment. Contract scope is typically locked at signature even though implementation is where the architecture reveals whether the original scope was calibrated correctly. Each of these patterns carries a plausible administrative rationale, which is part of why they endure. Outcome language protects the buyer, competitive scoring looks fair, and a fixed scope disciplines spending. Each also produces predictable friction once delivery begins, and that friction is rarely traced back to the procurement choice that caused it. The patterns are analytically distinct yet operationally entangled.

The perspective taken here is that of vendor-side solution architecture rather than buyer-side authority. Practitioner experience translating public-sector procurement vehicles, including the GSA Schedule, Texas DIR, HUB requirements, and the TIPS cooperative purchasing vehicle, into deliverable technical scope indicates that the translation gap is felt most acutely by the teams who convert contract language into working systems. This vantage point does not carry the authority of a procurement officer, and no such authority is claimed. It does offer close visibility into how RFP and RFI responses, statements of work, and pricing commitments interact with architectural reality after award. The vendor architect occupies the seam under discussion, reading a solicitation on one side and building against it on the other. That position affords a particular kind of evidence, anecdotal in form yet consistent across many engagements and vehicles.

The remainder of the paper proceeds in eight parts. Section 2 situates the argument within related work on AI in procurement and public sector adoption. Section 3 describes the three patterns of mistranslation, and Section 4 examines the structural causes that reproduce them. Section 5 proposes a reduction framework of three staged practices, ordered from earliest to latest intervention. Sector-specific applications across federal, state, and local vehicles follow in Section 6, where buyer capacity varies most. The paper then states its limitations in Section 7, discusses broader implications for AI governance in Section 8, and concludes in Section 9. Throughout, the claims are hedged deliberately, since the evidence is experiential rather than experimental.

2. Related Work

Institutional analysis of AI in public procurement has concentrated on modernizing the procurement process itself. A recent international review frames AI largely as a tool for improving core government functions, including procurement operations [6]. Guidance aimed at state procurement offices treats AI as a capability that can accelerate sourcing, evaluation, and market research [7]. Contract-level analysis documents a shift toward centralized buying and growing adoption of off-the-shelf AI products [8]. These sources describe, in useful detail, how the public sector buys AI and how that buying is changing. They give comparatively less attention to what happens after signature, when procurement language meets architecture during delivery. The process view and the delivery view rarely appear in the same frame. This paper works primarily in the second frame while borrowing context from the first.

A parallel body of public administration research examines how agencies adopt AI as an organizational undertaking. Implementation is described as a process shaped by capacity, culture, and strategic direction [9], and comparative case work shows adoption varying widely across otherwise similar organizations [10]. Adoption sometimes advances without authoritative strategic direction at all, driven instead by local initiative and individual champions [11]. Separate studies of competitive tendering consider where AI can assist the procurement function itself rather than the services being procured [12]. This literature illuminates adoption dynamics with care and empirical grounding. The specific seam between contract text and architectural decision, however, remains lightly treated within it. Adoption is studied as a whole, while the contractual joint that this paper isolates tends to disappear into the larger organizational story.

One strand of the literature comes closer to the present concern. Research on organizing public sector AI adoption frames a tension between separating and integrating specialized functions [13], which resonates directly with the split between procurement and architecture examined later. Broader analysis positions AI in the public sector as an opportunity paired with governance risk that must be actively managed [14], and conceptual frameworks propose smart governance structures for oversight [15]. Enterprise architecture research, for its part, treats flexibility as a precondition for digital transformation rather than a luxury [16], and systematic reviews describe how architecture practice is applied in e-government settings [17]. These threads

support, without directly naming, a procurement-to-architecture reading of failure. Drawing them together into one explanatory frame is part of the contribution attempted here.

Practitioner case studies of government AI deployments tend to attribute failure to vendor performance or to the limits of the technology. Empirical study of a labor market decision-support system, for instance, surfaces trust and design issues that are easy to read as vendor or model shortcomings [18]. Oversight of defense acquisitions similarly points to missing department-wide guidance rather than to a procurement-architecture seam [19]. These attributions are not wrong so much as incomplete, since they stop at the most visible actor in the chain. Reframing some of these failures as structural mistranslation may help future case studies locate root causes more accurately, one layer below the vendor. The reframing is offered as a complement to existing explanations rather than a replacement for them. Several causes can operate at once, and the structural one is the easiest to overlook.

Attention to organizational readiness rounds out the picture. Survey evidence from municipal managers indicates that adoption challenges genuinely matter for how public managers perceive and pursue AI [20]. Maturity assessment frameworks attempt to measure where an organization stands before it commits resources [21], and reviews of smart government catalog the recurring implementation challenges that agencies encounter [22]. Read together, this work suggests that the capacity to manage the procurement-architecture seam varies sharply across levels of government. A large federal agency and a small municipality face the same structural gap with very different resources. That variation, examined in Section 6, shapes which of the proposed practices a given buyer can realistically adopt, and how much of the gap it can close in practice.

3. Patterns of Mistranslation

The three patterns below share a common shape. Each begins with a reasonable procurement choice and ends with an architectural consequence that the choice did not anticipate. They are presented separately for analytical clarity, though they often appear together within a single engagement and can compound one another. An outcome-only specification, for example, is harder to correct when the scope is also locked. Table I summarizes each pattern alongside its mechanism, its typical failure mode, and a recommended fix. The recommendations are deliberately modest, since each is meant to fit inside existing procurement practice rather than to require its wholesale reform.

3.1 Outcome-Only Specifications

Solicitations commonly specify an outcome without the architecture that would make the outcome attainable. A requirement might state that a system will triage citizen service requests at ninety percent accuracy, offered here as an illustrative target rather than a figure drawn from any single deployment. The specification typically says nothing about data quality, the training corpus, integration with legacy systems, or the human review protocol. Those conditions largely determine whether the stated target is reachable, and many of them sit outside the vendor's direct control. A vendor can tune a model, but it cannot unilaterally clean a decade of inconsistent records or open a closed integration point. The requirement therefore transfers responsibility for an outcome without transferring authority over its preconditions. That imbalance between accountability and control is the defining feature of the pattern.

The dilemma this creates is visible most clearly from the vendor side of the table. A vendor that recognizes the missing conditions must either decline the opportunity or accept terms it cannot fully guarantee. A vendor that does not recognize them bids confidently and discovers the gap only during delivery. Price distortion follows, because the lowest bid often comes from the party that understood the preconditions least. The buyer, optimizing for price, may unintentionally select for misunderstanding. A practical remedy pairs each outcome specification with an explicit disclosure of the architectural conditions the buyer commits to maintain, such as data refresh cadence or integration availability. Naming those conditions converts a hidden risk into a shared and priced assumption, visible to every bidder at the same time.

3.2 Evaluation Without Architectural Test

Evaluation criteria reward what is easy to score before award. Proposals are graded on completeness, methodology, past performance, and price, and rarely on architectural feasibility within the buyer's actual environment [4]. The practical effect is to advantage teams skilled at proposal craft over teams skilled at delivery, since the two skills are not the same and are not held in equal measure. A polished response can describe, in confident prose, an approach that will not survive contact with the agency's data or its integration constraints. Because nothing in the evaluation tests that survival, the weakness stays invisible until implementation begins and the cost of discovery is high. The scoring rewards the wrong signal at precisely the moment the decision becomes irreversible. Good writing and good architecture are correlated only loosely.

A structured architectural review stage during evaluation would narrow this gap considerably. Shortlisted vendors could be asked to test their proposed approach against a representative slice of the buyer's environment, under controlled and equal conditions disclosed to all. Work on where AI assists the tendering process suggests that added analytical steps of this kind are feasible within competitive procedures [12]. The review need not be elaborate or expensive to be useful in practice. Even a modest feasibility check, applied consistently to every finalist, would surface architectural risk while the buyer still has genuine alternatives. Such a stage shifts selection toward vendors who can actually deliver rather than those who can merely describe. It also gives diligent vendors a place to demonstrate the rigor that outcome-only scoring otherwise ignores.

3.3 Locked Scope With Variable Architecture

Contracts fix scope at signature and treat any later change as an exception handled through amendment. For predictable work this discipline protects the buyer and keeps spending defensible. For AI the assumption strains, because implementation is exactly where the architecture reveals whether the scope was calibrated correctly in the first place. Enterprise architecture research treats flexibility as a precondition for digital transformation rather than a concession to be avoided [16]. A scope locked rigidly against that flexibility forces an awkward choice the moment reality diverges from the written assumptions. The amendment process is frequently too slow to absorb what delivery has just revealed, and the schedule keeps moving while the paperwork waits. The contract and the architecture drift apart precisely when they most need to stay aligned.

The consequence is a poor set of options for the delivery team. It can proceed with a scope it already knows to be wrong, or it can stall while an amendment moves through administrative review. Neither choice serves the public interest, and both waste effort that was contracted in good faith. Contract structures that anticipate architectural learning, with adjustment protocols operating within clearly defined bounds, offer a more workable third path. Bounded adjustment keeps accountability intact while still allowing the architecture to correct course as evidence accumulates. The aim is not open-ended change, which would reintroduce the very risks procurement rules guard against, but a disciplined channel for the learning that AI delivery reliably produces. The bounds are what make the flexibility defensible in the first place.

TABLE I. Three Patterns of Procurement-Architecture Mistranslation

Pattern	Mechanism	Typical Failure Mode	Recommended Fix
Outcome-only specifications	Outcome required without the architectural conditions it depends on	Vendors bid on preconditions outside their control; price distortion at award	Pair each outcome with disclosed conditions the buyer commits to maintain
Evaluation without architectural test	Scoring rewards proposal craft, not feasibility in the buyer's environment	Delivery-weak proposals win; architectural risk surfaces only in implementation	Add a structured architectural review stage against a representative environment
Locked scope with variable architecture	Scope fixed at signature; change handled as a slow amendment	Team proceeds with known-wrong scope or stalls awaiting an amendment	Use contract structures with bounded adjustment protocols for architectural learning

4. Structural Causes

Patterns that recur across many engagements usually have structural rather than individual causes. Blaming a particular officer or a particular vendor explains little when the same failure appears almost everywhere. Two structural features stand out in this domain. The first is the organizational separation of procurement and architecture into distinct functions. The second is an asymmetry inside vendor organizations between the teams that win work and the teams that deliver it. Table II sets out each cause, its organizational driver, and why it persists despite being widely understood. Both causes are ordinary features of how organizations divide labor, which is exactly why they prove so durable.

4.1 Procurement and Architecture as Separated Functions

In most public buyers, procurement and architecture are distinct functions with distinct owners and distinct accountabilities. Procurement officers own the contract language, the compliance obligations, and the vendor

relationship. Architecture leads own the technology decisions and the implementation oversight that follows award. The handoff between these two groups is rarely structured with much care. Procurement staff often lack deep architectural literacy, and architecture staff often lack procurement literacy or any real influence before signature. Research on organizing AI adoption frames precisely this tension between separated and integrated functions [13]. The arrangement is administratively tidy, because it gives each function a clear lane, yet it is operationally costly, because the AI-specific risks live in the space between the lanes where neither group holds full ownership.

The separation persists because it is reinforced from several directions at once. Reporting lines, performance metrics, and professional cultures all differ between the two functions, and the reporting chains often meet only at an executive level too senior for the working-level coordination the problem actually requires. Maturity assessments that measure organizational readiness rarely score this particular seam directly [21]. A modest countermeasure is to convene joint working sessions between procurement and architecture during requirement drafting and again before evaluation. Bringing both literacies into the same room early costs little beyond calendar time. It tends to catch mistranslations while they are still cheap to fix, well before they are frozen into a signed contract that only a slow amendment can later change.

4.2 Vendor-Side Asymmetry

A second cause sits inside the vendor organization rather than the buyer. Proposal teams optimize for winning the bid, while delivery teams inherit the resulting contract and must implement whatever it promised. The incentives facing these two groups are not symmetric. Proposal teams are seldom penalized for overpromising, since the shortfall surfaces long after award, and delivery teams are routinely penalized for underdelivering against commitments they did not set. This structure produces a predictable and systematic optimism about scope. From a solution architecture standpoint, the contract that finally reaches the delivery team frequently encodes assumptions that no architect ever endorsed. The gap is less dishonesty than a division of labor that rewards confidence at the bid stage and punishes it at the delivery stage.

A vendor-side remedy is to require architectural sign-off on proposals before they are submitted. Having an architect validate the scope, the data assumptions, and the integration claims adds friction to the bid process, and that friction is the entire point of the measure. It trades a slightly slower proposal for a substantially more deliverable one, which over time improves aggregate performance and the past-performance ratings on which later bids depend. The measure aligns the proposal team's incentive to win with the delivery team's obligation to perform, closing the internal loop. Comparative adoption evidence suggests such internal alignment remains uneven across organizations [10]. Firms that institutionalize the sign-off tend to protect their own delivery reputation, which is itself a competitive asset in repeat public procurement.

TABLE II. Structural Causes and Their Organizational Drivers

Cause	Organizational Driver	Why It Persists
Procurement and architecture separated	Distinct owners, literacies, and oversight roles	Separate reporting lines, metrics, and cultures; the chains meet only at executive level
Vendor-side asymmetry	Proposal teams win work; delivery teams inherit and perform it	Proposal teams are not penalized for overpromising; delivery teams are penalized for shortfalls

5. The Reduction Framework

The framework proposes three practices, each suited to a different point in the procurement lifecycle. They are ordered here from the earliest possible intervention to the latest, and a buyer should generally prefer the earliest that its rules and capacity allow. Table III compares the three on timing, applicability condition, and principal benefit. The practices are complementary rather than mutually exclusive, and a mature buyer may use all three across a portfolio of engagements. Each addresses the same underlying gap from a different distance, trading earlier influence for lower feasibility as the process advances toward award. In practice, the choice among them is dictated less by preference than by how much of the procurement timeline still remains when the gap is noticed.

5.1 Pre-Procurement Discovery

The earliest and usually most effective intervention happens before a formal solicitation exists at all. Structured discovery sessions with prospective vendors, conducted under appropriate ethical and transparency guidance, can explore architectural feasibility, data availability, integration constraints, and how a target outcome would actually be measured in production. Contract-level evidence shows the public sector still developing its collective understanding of how to buy AI well [8]. Discovery is where that understanding can be built collaboratively, before positions harden into a fixed specification. Its documented output then becomes the empirical foundation on which a better, more feasible solicitation is written. The buyer learns what is realistic, and the market learns what the buyer actually needs, before either side is committed to anything.

Discovery is especially valuable for novel technology categories where buyer and vendor understanding is still forming together. From the vendor side, an early technical discovery conversation is the same activity that precedes any well-scoped statement of work in commercial pre-sales. Bringing that familiar activity inside the public process, with equal access for all interested parties and carefully documented outputs, adapts a proven practice to a regulated setting. International guidance encourages exactly this kind of front-loaded engagement before requirements are fixed [6]. The result is a solicitation grounded in demonstrated feasibility rather than in aspiration. The discipline that keeps it fair is procedural, ensuring that no single vendor shapes the requirement in private while others are excluded from the same conversation.

5.2 Structured Architectural Briefings During Procurement

Where pre-procurement discovery is not possible, the next option operates inside the solicitation itself. The buyer can hold architectural briefings with all qualified bidders, or distribute documented written responses to every bidder, disclosing data conditions, integration points, governance constraints, and operational realities. A dedicated procurement standard for AI systems points toward exactly this kind of structured disclosure [4]. Transparency concerns, which are legitimate, are addressed by distributing identical information to all bidders at the same time. The briefing turns hidden architectural facts, which some vendors might infer and others miss, into shared inputs that every bid can price on equal footing. Fairness and feasibility, often treated as competing goals, are advanced together by the same act of disclosure.

This practice is lighter than full discovery and fits procedures that are already underway without disrupting them. It does not require the buyer to reopen a solicitation, only to enrich the information the solicitation already carries. The disclosures most worth making are precisely the ones vendors would otherwise learn only after award, at the point where correction is most expensive. Written responses distributed to all bidders preserve the procedural fairness that procurement rules exist to protect. The cost of preparing them is modest relative to the disputes they prevent. The information asymmetry a briefing removes is often the exact asymmetry that later produces change orders, schedule slippage, and the adversarial posture that follows a surprised delivery team.

5.3 Joint Architectural Reviews Post-Award

When neither earlier option is feasible, a meaningful intervention can still occur after award. The award itself can be made conditional on a joint architectural review within a defined window, commonly thirty to sixty days, producing a refined scope, price, and timeline before major work begins. Oversight of defense acquisitions has called for clearer guidance at exactly this stage of the acquisition lifecycle [19]. Within the review, the parties either adjust through agreed protocols or, if the originally awarded scope proves genuinely infeasible, terminate without penalty to either side. The window is deliberately short, long enough to test the critical assumptions yet brief enough to protect the overall schedule. It functions as a structured checkpoint rather than an open-ended renegotiation of the award.

A conditional review is simply more honest than proceeding with a scope both sides privately suspect is wrong. It gives the architecture a formal, contracted opportunity to inform the contract before large sums are irreversibly committed. From a delivery standpoint, the review is where a statement of work can be corrected while correction is still cheap and the working relationship is still cooperative. The option does carry a real cost in schedule and coordination, and it presumes good faith on both sides of the table. Where that presumption holds, it converts a likely downstream dispute into a planned and bounded checkpoint. The alternative, discovering the same facts six months later, is almost always more expensive and more acrimonious.

TABLE III. The Reduction Framework's Three Practices Compared

Practice	Timing	Applicability Condition	Key Benefit
Pre-procurement discovery	Before solicitation	Rules permit early vendor engagement	Solicitation grounded in feasibility, not aspiration
Structured architectural briefings	During procurement	Solicitation already underway	Removes information asymmetry every bid can price
Joint architectural review	Post-award (30-60 days)	Earlier options infeasible	Refines scope before large sums are committed

6. Sector-Specific Applications

The three practices do not apply uniformly, because procurement vehicles and buyer capacity differ sharply across levels of government. Federal, state, and local settings each impose distinct constraints and offer distinct room to maneuver. Table IV maps representative vehicles to their key constraints and the framework adaptation each suggests. The persona grounding in this section is direct, since these are the vehicles a vendor-side architect encounters repeatedly across state, higher-education, and cooperative engagements. What follows is not a compliance guide but a reading of where each practice is most likely to fit within the rules that already govern the vehicle.

Federal acquisition runs through vehicles such as GSA Schedules and IDIQ contracts, governed by a Federal Acquisition Regulation that predates AI as a procurement category. Adaptation has to respect small business set-asides, security clearance requirements, and a substantial burden of oversight reporting. Oversight bodies have noted that agencies still need to complete key AI requirements [5] and that defense acquisition in particular lacks department-wide guidance [19]. Within these constraints, structured briefings and conditional post-award reviews fit more easily than any wholesale redesign of the vehicle would. The regulation can accommodate architectural learning if the adjustment mechanisms are written directly into the task order rather than bolted on afterward. Working within the vehicle, rather than against it, is both faster and more defensible for a federal buyer.

State procurement relies heavily on cooperative purchasing and on state-specific frameworks that layer local policy onto the buy. The Texas DIR arrangement, with its Historically Underutilized Business participation requirements, illustrates how a state attaches additional policy goals to a technology procurement. Cross-state learning on these questions is being actively facilitated by state procurement associations [7]. From a vendor standpoint, translating a DIR or HUB-inflected solicitation into deliverable technical scope is a routine part of contract closure, and it is exactly where outcome-only specifications most often need architectural qualification. Pre-procurement discovery, where a state's rules permit it, tends to pay off most at this level, because the requirements are demanding yet the buyer usually retains enough capacity to run a genuine discovery process.

Local and municipal buyers face the tightest constraints of the three. Limited staff capacity, thin in-house architectural expertise, and smaller budgets together reduce their appetite for the upfront cost of a joint architectural review. Survey evidence that adoption challenges weigh heavily on public managers is consistent with this picture [20]. Cooperative vehicles such as TIPS and comparable regional cooperatives let local buyers borrow the frameworks and pre-negotiated terms developed at state or national scale. Leveraging an established cooperative is often far more realistic than building local AI procurement expertise from scratch, which few municipalities can justify for an occasional purchase. The framework's lightest practices, particularly the shared written briefing, are therefore the right starting point at this level, with heavier interventions reserved for larger buys.

TABLE IV. Sector-Specific Procurement Vehicle Considerations

Sector / Vehicle	Key Constraint	Framework Adaptation
Federal / GSA Schedule, IDIQ	FAR predates AI; set-asides, clearances, oversight reporting	Write briefings and conditional reviews into the task order

Sector / Vehicle	Key Constraint	Framework Adaptation
State / Texas DIR, HUB, cooperative	Layered policy goals and state-specific rules	Discovery where permitted; architectural qualification of outcome specs
Local / TIPS, regional cooperatives	Limited staff, expertise, and budget	Leverage cooperative frameworks; adopt the lightest practices first

7. Limitations

The argument rests on observed patterns rather than on a controlled empirical study. It generalizes from practitioner experience on the vendor side of public-sector engagements, which is at once a source of genuine insight and a source of predictable bias. The patterns described are plausible and, in this practitioner's experience, recurrent across vehicles, but they have not been measured against a systematic sample of engagements. Readers should therefore treat the claims as hedged propositions rather than as established empirical findings. The framework is best understood as a structured hypothesis offered for testing, not as a validated result. Its value, if any, lies in giving future empirical work a specific thing to look for.

Future research could test the framework by comparing agencies that adopt these practices with agencies that do not. Such a study would have to contend with selection effects, because agencies that choose to adopt may differ systematically from those that decline. It would also face sector confounds and the familiar difficulty of measuring public sector outcomes cleanly. A retrospective design across a set of completed engagements nonetheless appears feasible despite these obstacles, provided adoption can be coded reliably from contract records. Maturity assessment instruments could supply one axis of comparison for such a design [21]. The result would not be definitive, yet it would move the discussion from anecdote toward evidence, which is the direction this paper hopes to encourage.

The paper also addresses only one gap among several that afflict public sector AI. A procurement-to-operations gap, a gap between contracted and actual oversight capacity, and a gap between aspirational and operational equity commitments each deserve their own separate treatment. Isolating the procurement-to-architecture gap is a deliberate analytical choice, not a claim that it is the only gap that matters or even the largest. These gaps interact with one another in ways that would repay careful study, since a failure in one often propagates into the others. Treating them together in a single framework lies well beyond the present scope. The narrow focus is what allows the specific mechanisms here to be described with any precision.

A final limitation concerns good faith on both sides of the transaction. The framework presumes that buyer and seller negotiate honestly and intend to perform. A bad-faith buyer could extract concessions during discovery or review without offering any reciprocal commitment in return. A bad-faith seller could game the same discovery and review processes to shape a solicitation in its own favor. The framework's effectiveness therefore depends on a baseline of good faith, and additional governance mechanisms would be required wherever that baseline is absent. The practices are designed to reduce honest mistranslation between contract and architecture. They are not designed to deter deliberate manipulation, which calls for enforcement tools of a different kind.

8. Broader Implications

The translation gap is not unique to artificial intelligence. Cloud migration produced a strikingly similar gap during the 2010s, and the procurement community absorbed cloud-specific contract patterns only gradually, over several years of accumulated difficulty. AI appears likely to follow a comparable absorption path over the coming decade as the same lessons are relearned in a new context. Analysis of digital transformation in government suggests that these contractual adjustments tend to lag the underlying technology by years rather than months [2]. Documenting the pattern now, while the gap is still open, may modestly accelerate that absorption. Part of the value of the paper lies simply in naming a gap that many practitioners already feel but rarely articulate.

There is a public interest dimension to the argument as well. Public funds and public outcomes are at stake in every engagement, and better alignment between contract and architecture produces better public value for the same expenditure. Analysis of AI in the public sector frames the technology as an opportunity paired with risk that can be managed rather than merely feared [14]. High-volume procurement organizations stand to gain the most from these practices, because small per-engagement improvements compound across a large number

of buys. The argument is finally one of efficiency and stewardship rather than of novelty. Reducing avoidable failure is worthwhile even if the underlying mechanisms turn out to be unremarkable once they are named.

For AI governance specifically, procurement is a primary and still underused lever. It is one of the main mechanisms that determines how AI actually enters public service delivery, ahead of most downstream oversight. Conceptual frameworks for AI governance rightly emphasize oversight structures and accountability [15], yet the contract remains the place where those governance intentions become concrete operational constraints. Better procurement-architecture alignment is therefore a genuine leverage point for public sector AI governance as a whole, not merely a delivery convenience. Attention paid carefully at the contract stage propagates into every downstream decision about data, integration, and human oversight. Governing the contract is, in a real sense, governing the system that the contract produces.

The broader lesson generalizes well beyond any single procurement vehicle. Procurement language shapes implementation outcomes, and that relationship deserves sustained structural attention rather than case-by-case adjustment after each failure. Recognizing AI as a distinct procurement category, as recent guidance and standards have begun to do [3], [4], is a necessary first step but not a sufficient one. The structural view advanced here suggests that the logical next step is to redesign the seam between buying and building deliberately. That redesign lies within reach of practitioners working inside today's vehicles, using the modest practices described above. It does not require new legislation so much as new habits at the working level.

9. Conclusion

The procurement-to-architecture translation gap is a structural feature of public sector AI that the current literature has not addressed directly. The gap arises because procurement language, built for competition and accountability, does not by itself encode the architectural conditions on which AI outcomes ultimately depend. This paper has identified three patterns of mistranslation, traced them to two underlying structural causes, and proposed a reduction framework of three staged practices. Sector-specific considerations and candid limitations have been set out alongside the framework rather than after it. The account is diagnostic first and prescriptive second, on the view that a gap has to be named accurately before it can be closed.

The largest benefits come from joint adoption by both sides. A buyer that discloses architectural conditions and a vendor that validates scope before submission reinforce one another, and either practice alone helps meaningfully less than the pair. The framework is offered as practical guidance grounded in accumulated vendor-side experience rather than as a finished theory awaiting no revision. Practitioners can adopt elements of it today within existing vehicles, without waiting for regulatory change or new tooling. Researchers can refine and test it across diverse procurement contexts, and the resulting evidence would sharpen the guidance considerably. The gap will close as the community learns to write contracts that anticipate the architectures they commission.

References

[1] A. Cordella and N. Tempini, "E-Government and Organizational Change: Reappraising the Role of ICT and Bureaucracy in Public Service Delivery," Government Information Quarterly, vol. 32, no. 3, pp. 279-286, 2015. Available: https://doi.org/10.1016/j.giq.2015.03.005

[2] I. Mergel, N. Edelmann, and N. Haug, "Defining Digital Transformation: Results from Expert Interviews," Government Information Quarterly, vol. 36, no. 4, art. 101385, 2019. Available: https://doi.org/10.1016/j.giq.2019.06.002

[3] Office of Management and Budget, "Memorandum M-24-10: Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence," Executive Office of the President, Mar. 28, 2024. Available: https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf

[4] IEEE Standards Association, "IEEE Standard for the Procurement of Artificial Intelligence and Automated Decision Systems (IEEE 3119-2025)," 2025. Available: https://standards.ieee.org/ieee/3119/10729/

[5] U.S. Government Accountability Office, "Artificial Intelligence: Agencies Have Begun Implementation but Need to Complete Key Requirements," GAO-24-105980, Dec. 2023. Available: https://www.gao.gov/products/gao-24-105980

[6] OECD, "AI in Public Procurement," in Governing with Artificial Intelligence: The State of Play and Way Forward in Core Government Functions, OECD Publishing, Paris, 2025. Available:

https://www.oecd.org/en/publications/2025/06/governing-with-artificial-intelligence_398fa287/full-report/ai-in-public-procurement_2e095543.html
[7] NASPO, "AI in Procurement Is a Game Changer for Government," National Association of State Procurement Officials, Apr. 2025. Available: https://www.naspo.org/news/ai-in-procurement-is-a-game-changer-for-government/
[8] Kathrin Frauscher and Kaye Sklar, "The Surprising Shifts in How the Public Sector Is Buying AI, and What Policymakers Can Do About It," Open Contracting Partnership, Nov. 10, 2025. Available: https://www.open-contracting.org/2025/11/10/the-surprising-shifts-in-how-the-public-sector-is-buying-ai-and-what-policymakers-can-do-about-it/
[9] I. Mergel, H. Dickinson, J. Stenvall, and M. Gasco, "Implementing AI in the Public Sector," Public Management Review, vol. 27, no. 1, pp. 1-14, 2023. Available: https://doi.org/10.1080/14719037.2023.2231950
[10] O. Neumann, K. Guirguis, and R. Steiner, "Exploring Artificial Intelligence Adoption in Public Organizations: A Comparative Case Study," Public Management Review, vol. 26, no. 1, pp. 114-141, 2022. Available: https://doi.org/10.1080/14719037.2022.2048685
[11] A. Molin, "Examining Public Sector AI Adoption: Mechanisms for AI Adoption in the Absence of Authoritative Strategic Direction," in Proc. 25th Annual Int. Conf. Digital Government Research (dg.o 2024), Taipei, 2024, pp. 764-775. Available: https://doi.org/10.1145/3657054.3657278
[12] T. Hyppanen and J. Vilko, "AI in Public Procurement: Potential and Adoption in the Competitive Tendering Process," in Proc. 2025 IEEE Int. Conf. ICE/ITMC, Valencia, 2025, pp. 1-9. Available: https://doi.org/10.1109/ice/itmc65658.2025.11106613
[13] F. Selten and B. Klievink, "Organizing Public Sector AI Adoption: Navigating Between Separation and Integration," Government Information Quarterly, vol. 41, no. 1, art. 101885, 2024. Available: https://doi.org/10.1016/j.giq.2023.101885
[14] World Bank, Artificial Intelligence in the Public Sector: Maximizing Opportunities, Managing Risks, EFI Insight-Governance, Washington, DC: World Bank, 2020. Available: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/809611616042736565
[15] M. Rajagopal, R. Sivasakthivel, G. Ramar, Mansurali A., and S. K. Karuppasamy, "A Conceptual Framework for AI Governance in Public Administration: A Smart Governance Perspective," in Proc. 2023 7th Int. Conf. I-SMAC, Kirtipur, Nepal, 2023, pp. 488-495. Available: https://doi.org/10.1109/i-smac58438.2023.10290366
[16] Y. Gong, J. Yang, and X. Shi, "Towards a Comprehensive Understanding of Digital Transformation in Government: Analysis of Flexibility and Enterprise Architecture," Government Information Quarterly, vol. 37, no. 3, art. 101487, 2020. Available: https://doi.org/10.1016/j.giq.2020.101487
[17] D. K. W. W. Mahendra, S. Sulistyo, and G. D. Putra, "Enterprise Architecture in E-Government Application: A Systematic Literature Review," in Proc. 2024 Int. Electronics Symp. (IES), Denpasar, 2024, pp. 353-358. Available: https://doi.org/10.1109/ies63037.2024.10665873
[18] A. Berman, K. de Fine Licht, and V. Carlsson, "Trustworthy AI in the Public Sector: An Empirical Analysis of a Swedish Labor Market Decision-Support System," Technology in Society, vol. 76, art. 102471, 2024. Available: https://doi.org/10.1016/j.techsoc.2024.102471
[19] U.S. Government Accountability Office, "Artificial Intelligence: DOD Needs Department-Wide Guidance to Inform Acquisitions," GAO-23-105850, 2023. Available: https://www.gao.gov/products/gao-23-105850
[20] K. Sienkiewicz-Malyjurek, "Whether AI Adoption Challenges Matter for Public Managers? The Case of Polish Cities," Government Information Quarterly, vol. 40, no. 3, art. 101828, 2023. Available: https://doi.org/10.1016/j.giq.2023.101828
[21] M. Alfadhli, M. Kucukvar, N. C. Onat, S. Al-Maadeed, and A. Abdessadok, "Government Digital Transformation: A Tailor-Made Digital Maturity Assessment Framework," IEEE Access, vol. 13, pp. 71120-71132, 2025. Available: https://doi.org/10.1109/access.2025.3560999
[22] A. S. Prihatmanto, R. Andrian, W. D. Sunindyo, and R. Sutriadi, "Transforming Public Services: A Systematic Review of Smart Government Frameworks, Architectures, and Implementation Challenges," IEEE Access, vol. 12, pp. 135799-135810, 2024. Available: https://doi.org/10.1109/access.2024.3450907