

AI-Assisted Tender Normalization and Bid Evaluation for Construction Supply Chains in Saudi Arabia: A Data-Driven Procurement Framework for Vision 2030 Projects

Amaan Ullah Sharif Mohammed^{1*}
Independent Researcher

*Corresponding Author

Amaan Ullah Sharif Mohammed
Independent Researcher

Article History

Received: 19.08.2026

Accepted: 12.10.2026

Published: 14.10.2026

Abstract: Saudi Arabia's Vision 2030 construction programme is increasing the scale and complexity of tendering across infrastructure, buildings, tourism, transport and mega-project supply chains. Conventional bid evaluation remains vulnerable to inconsistent scope interpretation, fragmented tender documents, manual arithmetic checks, evaluator subjectivity and weak traceability between technical and commercial decisions. This review examines how artificial intelligence can support tender normalization and bid evaluation without displacing professional judgement or procurement accountability. Evidence published from 2020 to 2025 was synthesised across construction procurement, contractor selection, construction supply-chain digitalisation, natural language processing, explainable machine learning, supplier selection and AI-assisted bidding. The literature indicates that the principal opportunity is not autonomous award selection but the creation of a governed decision-support layer that converts heterogeneous tender submissions into a common comparison basis. Natural language processing can extract qualifications, exclusions, scope assumptions, programme commitments and contractual risks; rules and structured analytics can reconcile quantities, currencies, unit rates, provisional sums and arithmetic differences; and explainable models can support technical, commercial, sustainability and supplier-risk scoring. However, automation also creates risks of hidden bias, data leakage, inconsistent criteria, overreliance, algorithmic collusion and weak legal defensibility. This review therefore proposes a six-stage framework comprising data ingestion, document intelligence, tender normalization, multi-criteria evaluation, human governance and award recommendation, with project-performance feedback used to recalibrate future models. For Saudi Vision 2030 projects, the framework offers a route to faster, auditable procurement at scale while preserving fairness, local-content considerations, competition and accountable human approval.

Keywords: Artificial Intelligence, Tender Normalization, Bid Evaluation, Construction Procurement, Construction Supply Chain, Natural Language Processing, Contractor Selection, Saudi Arabia, Vision 2030, explainable AI.

Copyright © 2026 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY-NC 4.0) which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

Citation: Mohammed, A. U. S. (2026). AI-Assisted Tender Normalization and Bid Evaluation for Construction Supply Chains in Saudi Arabia: A Data-Driven Procurement Framework for Vision 2030 Projects. *Glob Acad J Econ Buss*, 8(5), 1818-1828. <https://doi.org/10.36348/gajeb.2026.v08i05.067>

1. INTRODUCTION

Construction tender evaluation is one of the most consequential decision points in the project supply chain. The award decision fixes not only the immediate contract price but also the contractor, subcontractor and supplier capabilities that shape programme reliability, quality, safety, claims exposure, cash flow and downstream supply continuity. In Saudi Arabia, the significance of this decision is amplified by the scale of Vision 2030 development, where major programmes require repeated procurement of complex packages across infrastructure, real estate, transport, utilities, tourism and industrial projects. Digital procurement is therefore becoming less a convenience than an institutional capability for controlling complexity, improving transparency and protecting value.

Research on construction e-procurement and digital supply chains has already established that electronic platforms can improve information exchange, process integration and sustainability when organisational and governance barriers are addressed [1-8]. More recent construction-supply-chain literature shows that data fragmentation, coordination failures, disruption risk and limited integration remain persistent challenges, while AI and other digital technologies are increasingly proposed as mechanisms for improving resilience and decision quality [13-23]. Saudi-focused studies similarly indicate that digital transformation, process automation and supplier governance can improve procurement efficiency, but readiness, interoperability and organisational capability are uneven [9-30].

Tender evaluation is particularly difficult because bids are rarely directly comparable when first received. One bidder may exclude temporary works that another includes; another may price provisional sums differently, use an alternative specification, assume a different programme, qualify warranty obligations, omit testing or commissioning, or distribute cost differently across a bill of quantities. Commercial totals can therefore conceal material differences in scope and risk. The practical process known as tender normalization, bid leveling or bid equalisation attempts to convert these heterogeneous submissions to a consistent basis before scoring and award. In large packages, the task is labour intensive because evaluators must reconcile structured price schedules with unstructured technical proposals, clarifications, qualifications, exclusions, method statements, programmes and contractual departures.

AI is well suited to parts of this problem because contemporary models can process both structured and unstructured information. In

construction, machine learning is increasingly used across the project lifecycle, while natural language processing is being applied to contracts, procurement documents and risk classification [17-25]. Bid-evaluation research has also advanced beyond lowest-price selection toward multi-criteria and uncertainty-aware methods that combine price, technical capability, experience and other attributes [2-15]. Explainable supplier-selection studies further show that machine learning can be integrated with familiar multi-criteria decision-making rather than replacing it [19]. These developments suggest a practical design principle: AI should automate evidence extraction, normalization and analytical consistency while human evaluators retain authority over criteria, interpretation, exceptions and award approval.

The challenge is governance. Algorithmic assistance can amplify existing bias if historical awards are treated as unquestioned training labels. Models can mistake linguistic style for substantive quality, underweight innovative alternatives, or conceal how a ranking was produced. AI-based bidding systems may even change competition itself; simulation studies demonstrate the possibility of algorithmic coordination and collusive pricing under repeated bidding environments [20-26]. Consequently, an AI-assisted tender process must be designed around traceability, contestability and human oversight rather than predictive accuracy alone.

This review develops such a framework for construction supply chains in Saudi Arabia. It integrates evidence on digital procurement, bid evaluation, contractor and supplier selection, NLP, construction AI, procurement risk and algorithmic bidding. The central argument is that tender normalization should be treated as a governed data transformation process: raw bids are converted into a common commercial and technical basis, every adjustment is evidenced, the evaluation criteria remain pre-defined, and AI outputs are presented as reviewable decision support rather than autonomous determinations.

2. Aim, Objectives and Review Questions

The aim of this review is to develop an evidence-informed framework for AI-assisted tender normalization and bid evaluation suitable for construction supply chains supporting Saudi Vision 2030 projects.

Four objectives guide the review. First, it identifies the information problems that make construction bids difficult to compare, including inconsistent scope, pricing structures, qualifications and document formats. Second, it evaluates how NLP,

machine learning and rule-based analytics can support document extraction, normalization and contractor or supplier assessment. Third, it examines governance requirements for explainability, fairness, anti-collusion, auditability and human challenge. Fourth, it translates the evidence into an implementation framework that links tender data to award decisions and post-award supplier performance.

Three review questions are addressed: RQ1, which tender information can be reliably standardized or extracted by AI, and which elements still require professional interpretation? RQ2, how should normalized commercial data be combined with technical, capability, sustainability and risk criteria without obscuring evaluator judgement? RQ3, what governance controls are required to make AI-assisted bid evaluation transparent, reproducible and defensible in high-value Saudi construction procurement?

3. REVIEW METHODOLOGY

A structured integrative review was adopted because the topic cuts across construction management, procurement, supply-chain management, artificial intelligence, NLP and decision science. The evidence window was restricted to 2020-2025. Priority was given to peer-reviewed studies that addressed construction e-procurement, bid or contractor evaluation, construction supply-chain digitalisation, automated document analysis, explainable supplier selection, AI-assisted bidding, digital procurement or Saudi construction-sector transformation. International studies were retained where the mechanism was transferable to Saudi projects.

Search concepts combined terms relating to construction tendering, bid evaluation, tender normalization, contractor selection, supplier selection, procurement, artificial intelligence, machine learning, natural language processing, large language models, explainability, construction supply chain, Saudi Arabia and Vision 2030. Evidence was screened for direct relevance to at least one of six analytical themes: tender data preparation; normalization; evaluation criteria; AI/NLP methods; governance and fairness; or implementation and feedback. Studies focused exclusively on autonomous prediction without a clear procurement use case were treated as supporting rather than central evidence.

The final synthesis used 30 DOI-resolvable references. Deductive coding grouped findings under data ingestion, document intelligence, normalization, scoring, governance and post-award learning. Inductive coding then identified recurring

operational mechanisms, including requirement extraction, scope mapping, arithmetic reconciliation, currency and unit conversion, exclusion detection, clause classification, qualification tracking, supplier-risk assessment, evaluator consensus, explanation, override logging and performance feedback. No primary data or invented effect sizes are reported. The proposed framework is therefore a conceptual synthesis of recent evidence rather than an empirical claim that AI has already produced specified savings or award outcomes in Saudi projects.

4. Why Construction Tenders Require Normalization

The difficulty of tender comparison begins with the fact that a tender package is a bundle of related but differently structured information. Bills of quantities and commercial schedules may be tabular, while technical proposals, clarifications, qualifications and method statements are narrative. Drawings and specifications define scope indirectly through references and interfaces. Contract conditions distribute risks that may later appear as premiums, exclusions or alternative offers. A simple comparison of submitted totals can therefore produce a false sense of equivalence.

Traditional bid-evaluation research recognises this multidimensionality. Multi-criteria models have been developed to combine price and non-price attributes under uncertainty, including expertise-weighted and linguistic decision frameworks [2-4]. Contractor-selection reviews show that financial capacity, technical capability, experience, management, quality and safety commonly coexist with price in prequalification and award decisions [7]. Studies of road and renovation procurement likewise demonstrate that innovation, risk and qualitative characteristics can materially influence the selection result [5-11]. The practical implication is that normalization must occur before aggregation: inputs need a common meaning before scores can be combined.

Commercial normalization typically addresses five categories. The first is arithmetic consistency: extensions, subtotals, taxes, discounts and unit-rate calculations must reconcile. The second is scope consistency: exclusions or qualifications are converted into transparent adjustments or evaluation flags. The third is measurement consistency: units, currencies, price bases and escalation assumptions are aligned. The fourth is risk consistency: provisional sums, contingencies, alternative materials, warranties and programme assumptions are compared on an equivalent basis. The fifth is time consistency: payment terms, delivery dates and programme commitments are interpreted against a shared baseline.

AI can assist because these tasks mix deterministic checks with semantic interpretation. Arithmetic reconciliation and currency conversion are rule based. Scope differences, contractual departures and ambiguous qualifications require textual analysis. Construction procurement repositories already contain large amounts of semi-

structured data, but data-access and preparation remain barriers to machine-learning adoption [16]. The first design requirement is therefore a data layer that can ingest spreadsheets, PDFs, schedules and narrative documents while preserving source provenance.

Table 1: Tender normalization problems and corresponding AI-assisted controls

Problem area	Typical bid variation	AI-assisted treatment	Mandatory human control
Commercial arithmetic	Formula errors, missing extensions, inconsistent subtotals	Automated recalculation, cross-foot checks, anomaly flags	Validate commercial rules and approved corrections
Scope coverage	Inclusions, exclusions, optional items, interface gaps	Clause extraction, semantic matching, compliance matrix	Confirm interpretation and materiality of scope difference
Units and currencies	Mixed currencies, units, exchange bases	Structured conversion and unit mapping	Approve reference date, rate source and contractual basis
Qualifications	Technical departures, assumptions, warranties, programme caveats	NLP classification and similarity mapping	Determine acceptability and risk transfer
Provisional and alternative pricing	Provisional sums, options, substitutions	Normalization scenarios and sensitivity analysis	Decide comparable basis and treatment of alternatives
Supplier capability	Experience, resources, financial and delivery evidence	Feature extraction and risk scoring	Challenge model output and confirm evidence quality

5. AI-Assisted Document Intelligence

Document intelligence is the bridge between tender files and a normalized comparison. Construction procurement generates large volumes of text, and recent evidence shows that NLP can transform this material into structured signals. Automated contract analysis has classified construction clauses by obligation, right, risk and responsible party, demonstrating that machine learning can support faster review of lengthy contractual material [25]. In procurement more broadly, transformer models can identify subtle contextual patterns in complex tender text, while file-scraping research shows that bills of quantities and specifications can be systematically converted into machine-readable datasets for downstream NLP and machine learning [16].

For tender evaluation, the most useful NLP tasks are not unrestricted summarisation but constrained extraction and classification. A requirements extractor can identify mandatory criteria, requested deliverables, performance specifications and evidence obligations. A bid parser can then locate the corresponding response, classify it as compliant, partially compliant, non-compliant or unclear, and link the judgement to the original passage. Similar methods can extract exclusions, proposed alternatives, subcontracting arrangements, programme commitments and contractual qualifications.

Long documents require special treatment. Tender specifications can exceed the context length of conventional models, and important clauses may be dispersed across appendices. The system should therefore preserve document hierarchy, section identifiers and cross-references rather than flattening all text into one sequence. Retrieval should return the source clause with each extracted output so an evaluator can verify meaning. This requirement is more important than model elegance: a high-scoring model that cannot show the evidential basis of a deviation is unsuitable for defensible procurement.

Multilingual capability is also relevant in Saudi Arabia. Major projects may contain English and Arabic content, while supplier material can vary in terminology and writing style. Cross-lingual models provide a technical route to multilingual processing, but translation or cross-lingual embeddings should not silently change contractual meaning. High-risk terms, quantities, dates, guarantees and legal qualifications require source-language verification. AI should therefore identify and organize evidence, not become the final interpreter of contractual obligations.

6. Normalization as a Governed Data Transformation

Normalization should be treated as an explicit transformation pipeline rather than an informal spreadsheet exercise. Every raw value or

clause should have a traceable path to the normalized output. This can be achieved through four linked objects: the original bid item, the extracted interpretation, the normalization rule, and the approved adjustment or flag.

The first normalization layer is structural. Different bidders may use different schedule layouts or item descriptions. Semantic matching can align comparable items across bids and map them to a reference work breakdown structure. Structured price data can then be reconciled against quantities and unit rates. Where confidence is low, the system should generate a clarification task rather than force a match.

The second layer is commercial. Arithmetic corrections, discount logic, tax treatment, currency conversion and unit normalization can be automated using deterministic rules. The system should retain both the submitted and normalized value. This distinction is essential because normalization is an analytical construct, not permission to rewrite a supplier's legal offer.

The third layer concerns scope and risk. If one bidder excludes commissioning while others include it, the engine can identify the exclusion, map it to the relevant scope requirement and estimate a comparison adjustment if an approved benchmark exists. Where no reliable benchmark exists, it should flag the issue for evaluator judgement. The same principle applies to alternative products, programme assumptions, access restrictions, temporary works and warranty departures.

The fourth layer is uncertainty. Normalization does not remove all ambiguity, and a single adjusted total may conceal the sensitivity of the ranking. Scenario analysis should therefore show how bidder positions change under plausible treatments of unresolved items. This approach is consistent with decision research that treats construction bid evaluation as an uncertainty problem rather than a deterministic ranking exercise [2-11]. It also reduces pressure to encode contested assumptions as if they were facts.

7. Multi-Criteria Bid Evaluation after Normalization

Once bids have been normalized to a common basis, evaluation can combine commercial and non-commercial dimensions. The literature supports multi-criteria approaches because

contractor and supplier selection is inherently a balancing problem [2-7]. Machine learning can improve the evidence base, but the criteria architecture must remain intelligible to procurement stakeholders.

A suitable evaluation hierarchy for large Saudi construction packages has five domains. Mandatory compliance determines whether a bid is eligible for weighted scoring. Technical capability examines methodology, design response, programme, resources and quality. Commercial evaluation uses normalized price, lifecycle implications and material financial qualifications. Supplier capability considers financial resilience, delivery record, capacity, subcontracting and supply-chain exposure. Strategic criteria can include sustainability and Saudi local-content requirements where they are part of the published tender basis.

Explainable supplier-selection research shows the value of retaining familiar multi-criteria decision structures while using machine learning to reduce complexity and improve objectivity [19]. Intelligent criteria-construction studies similarly demonstrate that data can help identify which supplier attributes are most informative [14]. For construction procurement, this suggests a hybrid design: AI extracts and structures evidence, highlights patterns and tests sensitivity, while the evaluation committee approves criteria, weights and final scores.

AI can also identify abnormal scoring patterns. If one evaluator consistently scores a bidder materially above the panel, the system can flag the divergence for moderation without assuming that the minority score is wrong. It can compare narrative comments with numeric scores, identify missing justifications, or detect that similar evidence has been treated differently across bidders. Such functions improve consistency while preserving evaluator independence.

The commercial score should never be based on raw submitted price where material scope differences remain. Instead, the evaluation should show submitted price, approved normalization adjustments, normalized comparison price and unresolved risk separately. This separation allows management to see whether the apparent lowest bidder remains competitive after scope equalisation and whether the ranking is sensitive to assumptions.

Table 2: Proposed evaluation architecture for Vision 2030 construction tenders

Evaluation domain	Illustrative evidence	AI support	Governance requirement
Mandatory compliance	Licences, bonds, required forms, minimum experience	Document completeness and pass/fail extraction	No model override of explicit tender rules
Technical response	Methodology, design compliance, programme, resources	Requirement matching, evidence extraction, consistency checks	Pre-defined rubric and evaluator review
Commercial value	Normalized price, payment, exclusions, lifecycle cost	Arithmetic reconciliation, normalization, scenario analysis	Approved normalization rules and visible adjustments
Supplier capability	Delivery history, financial resilience, capacity, claims history	Risk features, trend analysis, supplier segmentation	Verify data relevance and avoid historical-bias lock-in
Strategic criteria	Sustainability, local content, innovation where specified	Evidence classification and score support	Criteria disclosed before tender and applied consistently
Award assurance	Consolidated score, exceptions, sensitivities, unresolved risks	Explanation package and ranking scenarios	Documented moderation, approval authority and audit trail

8. Supplier and Construction Supply-Chain Intelligence

Tender evaluation should not end at the bidder's proposal. Construction supply-chain research shows that project outcomes depend on networks of suppliers, subcontractors, logistics flows and information relationships [13-18]. A bidder may present a strong headline offer while relying on weak or concentrated downstream suppliers. AI-assisted evaluation can therefore incorporate evidence on supply-chain resilience where this information is legitimately available and relevant to the tender criteria.

Digital procurement studies demonstrate increasing interest in connected supply-chain technologies, while Saudi construction research has examined blockchain readiness as a means of improving traceability and resilience [8-10]. The practical lesson is broader than any single technology: supplier information should become reusable institutional data rather than disappearing after award. Delivery reliability, variation history, quality performance, claims patterns, responsiveness and approved-supplier status can form a controlled performance record for future procurement.

However, historical performance data can introduce bias. Incumbent suppliers naturally have richer records than new entrants, and past project conditions may not be comparable. Performance features should therefore be normalised for context and separated from subjective reputation. A model should not simply learn that previous winners are safer choices. Procurement governance must preserve competitive entry and allow bidders to challenge inaccurate historical data.

Supplier-selection studies using machine learning and hybrid optimization illustrate how data-driven methods can support complex multi-criteria decisions [14-29]. Their relevance to construction lies not in copying their models directly but in adopting the principle of transparent feature selection, uncertainty testing and sensitivity analysis. The goal is to identify decision-relevant evidence, not to create an opaque supplier score that substitutes for due diligence.

9. Governance, Explainability and Fairness

Governance is the condition that makes AI-assisted evaluation usable in high-value procurement. The framework must be capable of explaining not only the final ranking but also how data moved from the supplier's submission to the normalized value and score. This requires data lineage, version control, rule ownership, model documentation, role-based access, override logging and retained evidence.

Explainability should be role specific. Commercial evaluators need to see how prices were adjusted and which bid clauses caused the adjustment. Technical evaluators need requirement-to-response traceability. Procurement leaders need sensitivity analysis and unresolved risks. Auditors need a reproducible record of source files, model version, evaluator changes and approvals. Suppliers, where procedures permit disclosure, need an intelligible basis for understanding the evaluation outcome.

Fairness requires identical analytical treatment across bidders. The same extraction logic, normalization rule and scoring rubric should apply to all submissions. Models should be tested for

systematic differences associated with writing style, document length, incumbent status or other attributes that are unrelated to procurement merit. Human overrides are legitimate, but the original output and reason for change should remain visible.

Competition risk also needs explicit monitoring. AI-assisted bid pricing can change market behaviour, and construction studies have shown that repeated algorithmic interaction may facilitate collusive outcomes [20-28]. The buyer's AI system should therefore avoid functions that inadvertently reveal competitor-sensitive information or train bidders toward common pricing behaviour. Anti-collusion analytics can monitor suspicious similarity, repeated bidder relationships or unusual pricing patterns, but such signals should initiate review rather than constitute proof.

Cybersecurity and confidentiality are equally important. Tender submissions contain commercially sensitive pricing, methodologies and intellectual property. Training or evaluating third-party models on live bid data without approved controls can expose information beyond the procurement environment. Sensitive tenders therefore require controlled deployment, access restrictions, encryption, data-retention rules and contractual clarity regarding model providers.

Figure 1 presents the proposed end-to-end architecture, while Figure 2 shows the closed-loop governance model. Together they emphasise that AI sits inside the procurement control environment rather than outside it.

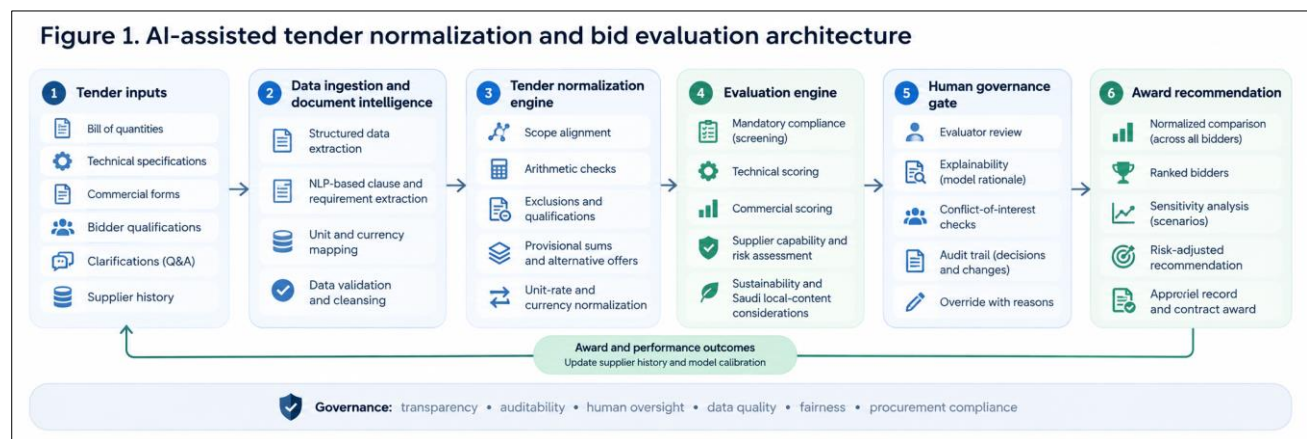


Figure 1: AI-assisted tender normalization and bid evaluation architecture

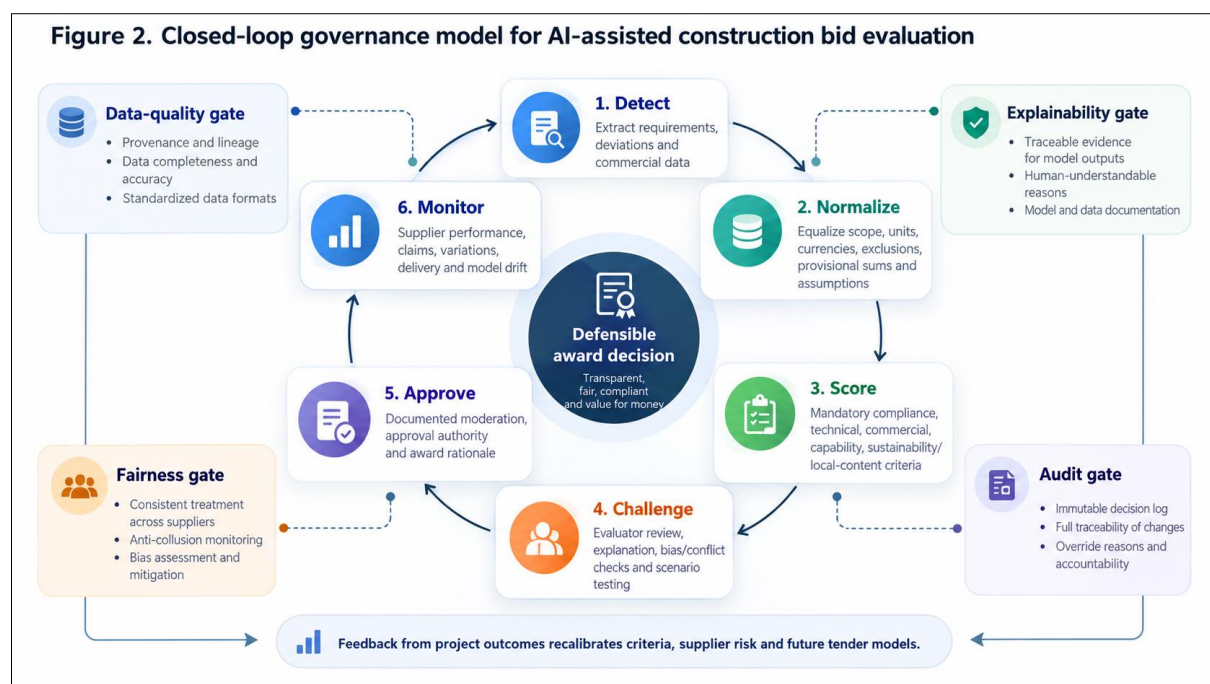


Figure 2: Closed-loop governance model for AI-assisted construction bid evaluation

A further implementation issue is the separation between evidence extraction and evaluative inference. Procurement teams should not allow a language model to move directly from raw tender text to an award score. A safer design uses intermediate structured records that can be reviewed independently: requirement identifier, bidder response location, extracted value, confidence, normalization rule, evaluator decision and final treatment. This creates an evidence ledger that can be reconstructed after award and protects the process when staff change or a decision is challenged. The ledger also enables model validation because extraction errors can be distinguished from scoring disagreements.

Data architecture should support versioned tender documents, addenda and clarifications. If a specification changes after tender release, the system must know which requirement version each bidder answered and whether a clarification modifies the commercial basis. Similarly, post-tender clarifications should be stored as new evidence rather than overwriting original submissions. This matters in major projects where multiple clarification rounds can otherwise obscure the chronology of decision-making.

Model performance should therefore be measured at task level. Extraction accuracy, missed qualification rate, false compliance rate, arithmetic exception detection, mapping confidence and evaluator override frequency are more useful than a single overall accuracy score. Procurement leaders can then identify where automation is reliable and where human review must remain intensive. In addition, periodic blind re-performance by independent reviewers can test whether AI assistance improves consistency without anchoring evaluators to machine suggestions. Such controls make the technology auditable, allow proportional use by risk level, and provide a defensible pathway for scaling from low-consequence document checks to higher-consequence evaluation support. They also create measurable acceptance criteria for procurement teams, auditors, legal reviewers and executive approvers.

10. Vision 2030 Implementation Framework

The proposed framework has six operational stages. Stage one is tender data ingestion. Procurement teams define the reference tender structure, mandatory criteria and controlled data sources before bids are received. Submitted documents are preserved in original form and assigned immutable identifiers.

Stage two is document intelligence. Structured schedules are parsed, while NLP extracts

requirements, qualifications, exclusions, alternative offers and commitments. Every extracted item links back to a source location and receives a confidence or review status.

Stage three is tender normalization. Deterministic rules correct arithmetic and align units, currencies and price bases. Semantic mapping aligns scope items and qualifications. Approved adjustments are recorded without replacing the original bid, and unresolved issues become clarification or risk items.

Stage four is evaluation. Mandatory screening precedes weighted scoring. Technical, commercial, supplier-capability and strategic criteria are evaluated against pre-defined rubrics. AI can provide evidence summaries, anomaly detection and scenario comparison, but evaluators remain responsible for the score.

Stage five is governance and approval. Evaluators review explanations, resolve conflicts, perform moderation and record reasons for overrides. Sensitivity analysis tests whether the recommended bidder remains preferred under plausible normalization assumptions. The award recommendation therefore includes both the ranking and the unresolved risks.

Stage six is post-award learning. Supplier and project outcomes are measured against tender commitments. Delivery reliability, variations, claims, quality, programme performance and supply-chain disruptions are fed back into controlled supplier records and model validation. This creates a learning system rather than a one-off evaluation tool.

Implementation should begin with narrow, high-volume tasks that are easy to verify. Suitable pilots include arithmetic checks, completeness screening, clause extraction, exclusion identification and bid-tabulation support. More consequential functions such as technical scoring or supplier-risk prediction should follow only after data quality, evaluator trust and governance controls have matured [27]. EPC research similarly suggests that AI adoption should progress through business-value definition, data readiness, deployment choices and organisational capability rather than technology-first experimentation [21].

Saudi deployment also requires local context. Tender documents may be bilingual, local-content requirements may be package specific, and public, PIF-related and private procurement environments may differ in authority and evaluation rules. The framework should therefore be configurable rather than hard coded. It should

support project-specific scoring policies while enforcing common controls around evidence, auditability and human approval.

11. Discussion and Research Gaps

The reviewed literature supports the feasibility of individual components but reveals a gap at their intersection. Construction studies address bid evaluation, contractor selection, AI, contract NLP and supply-chain digitalisation, yet relatively little research integrates tender normalization, document intelligence and governed award decision-making into one architecture. This gap is particularly important because normalization is where many practical procurement errors arise.

A second gap concerns ground truth. Historical award decisions are poor labels for supervised learning if the objective is to improve future evaluation. Past awards may reflect outdated criteria, incomplete information or institutional bias. Better training labels could be built from post-award outcomes, verified compliance findings, claims performance and expert-reviewed normalization records. This would move model learning from "predict the previous winner" toward "identify evidence associated with successful delivery and defensible evaluation."

A third gap concerns explainability quality. Many AI studies measure predictive performance more rigorously than whether explanations help procurement professionals make better decisions. Future research should test whether evaluators can reconstruct a model's reasoning, identify errors, challenge inappropriate recommendations and reach more consistent outcomes. Explanation usefulness should be treated as an operational metric rather than a visual feature.

A fourth gap is Saudi-specific validation. Local datasets are needed to test bilingual extraction, local-content evidence, contractor categories, project-package differences and supply-chain structures. The expanding portfolio of Vision 2030 projects offers a strong environment for comparative research, but data governance must protect commercial confidentiality. Controlled multicentre studies could compare manual and AI-assisted evaluation on time, error rates, consistency, clarification volume and post-award outcomes.

Finally, the market impact of procurement AI requires continued attention. AI can improve buyer-side consistency while supplier-side algorithms simultaneously optimize pricing and strategy. Research on algorithmic collusion demonstrates that evaluation method and repeated interactions can influence bidding behaviour [20-26]. Future

frameworks should therefore evaluate not only internal efficiency but also competition, entry, diversity and long-term market effects.

12. CONCLUSION

AI-assisted tender normalization has the potential to improve one of the most labour-intensive and judgement-sensitive stages of construction procurement. Its value lies less in replacing evaluation committees than in creating a consistent evidence layer: documents are parsed, scope differences exposed, commercial data reconciled, criteria applied transparently and every adjustment linked to its source.

The evidence suggests that construction procurement is ready for a hybrid model in which deterministic controls, NLP, machine learning and multi-criteria evaluation operate together. Tender normalization should precede weighted scoring; explainable analytics should support rather than obscure professional judgement; and supplier intelligence should be linked to verified project outcomes rather than reputation alone. Human challenge, audit trails, anti-collusion safeguards and data governance are essential because high predictive performance cannot compensate for an indefensible award process.

For Saudi Arabia, this approach aligns directly with the need to deliver a large portfolio of Vision 2030 projects through resilient and transparent construction supply chains. A staged implementation beginning with document extraction and normalization can build the data quality and institutional trust required for more advanced evaluation. The proposed six-stage framework provides a practical route from heterogeneous tender submissions to a normalized, explainable and risk-aware award recommendation, while post-award feedback creates continuous improvement. The strategic objective is therefore not autonomous procurement, but faster and more reliable human decision-making supported by traceable data, consistent rules and accountable AI.

REFERENCES

1. Yu, A.T.W.; Yevu, S.K.; Nani, G. Towards an integration framework for promoting electronic procurement and sustainable procurement in the construction industry: A systematic literature review. *Journal of Cleaner Production* 2020, 250, 119493. <https://doi.org/10.1016/j.jclepro.2019.119493>
2. Chen, Z.-S.; Zhang, X.; Pedrycz, W.; Wang, X.-J.; Skibniewski, M.J. Bid evaluation in civil construction under uncertainty: A two-stage LSP-ELECTRE III-based approach. *Engineering Applications of Artificial Intelligence* 2020, 94,

103835.
<https://doi.org/10.1016/j.engappai.2020.103835>
3. Chen, Z.-S.; Zhang, X.; et al. Bid Evaluation for Major Construction Projects Under Large-Scale Group Decision-Making Environment and Characterized Expertise Levels. *International Journal of Computational Intelligence Systems* 2020, 13, 1221-1242.
<https://doi.org/10.2991/ijcis.d.200801.002>
4. Chen, Z.-S.; Zhang, X.; Rodríguez, R.M.; Pedrycz, W.; Martínez, L. Expertise-based bid evaluation for construction-contractor selection with generalized comparative linguistic ELECTRE III. *Automation in Construction* 2021, 125, 103578.
<https://doi.org/10.1016/j.autcon.2021.103578>
5. Oad, P.K.; Kajewski, S.; Kumar, A.; Xia, B. Bid Evaluation and Assessment of Innovation in Road Construction Industry: A Systematic Literature Review. *Civil Engineering Journal* 2021, 7, 179-196.
<https://doi.org/10.28991/cej-2021-03091646>
6. Oad, P.K.; Kajewski, S.; Kumar, A.; Xia, B. Investigation of Innovation during Bid Evaluation Process in the Road Construction Industry. *Civil Engineering Journal* 2021, 7.
<https://doi.org/10.28991/cej-2021-03091676>
7. Acheamfour, V.K.; Kissi, E.; Adjei-Kumi, T.; Adinyira, E. Review of empirical arguments on contractor selection models for construction projects. *International Journal of Construction Management* 2021.
<https://doi.org/10.1080/15623599.2020.1868092>
8. Yevu, S.K.; Yu, A.T.W.; Darko, A. Digitalization of construction supply chain and procurement in the built environment: Emerging technologies and opportunities for sustainable processes. *Journal of Cleaner Production* 2021, 322, 129093.
<https://doi.org/10.1016/j.jclepro.2021.129093>
9. Alabdali, M.A.; Salam, M.A. The Impact of Digital Transformation on Supply Chain Procurement for Creating Competitive Advantage: An Empirical Study. *Sustainability* 2022, 14, 12269.
<https://doi.org/10.3390/su141912269>
10. Al Azmi, N.; Sweis, G.; Sweis, R.; Sammour, F. Exploring Implementation of Blockchain for the Supply Chain Resilience and Sustainability of the Construction Industry in Saudi Arabia. *Sustainability* 2022, 14, 6427.
<https://doi.org/10.3390/su14116427>
11. Li, W.; Li, Q.; Liu, Y.; Wang, S.; Pei, X. A new method of bid evaluation for renovation projects: Based on unascertained measure theory and entropy weight. *PLoS ONE* 2022, 17, e0271534.
<https://doi.org/10.1371/journal.pone.0271534>
12. Zhang, Y.; Chen, Q. Identifying Key Influential Factors of Bid Evaluation in Government Public Project Green Procurement in China Using BP-DEMATEL Model. *Mathematical Problems in Engineering* 2022, 2022, 8223757.
<https://doi.org/10.1155/2022/8223757>
13. Liao, L.; Yang, C.; Quan, L. Construction supply chain management: A systematic literature review and future development. *Journal of Cleaner Production* 2023, 382, 135230.
<https://doi.org/10.1016/j.jclepro.2022.135230>
14. Wu, S.; Jia, F.; Barnes, D. Criteria system construction for sustainable supplier selection: A novel approach based on a matching model and machine learning. *Applied Soft Computing* 2023, 148, 110586.
<https://doi.org/10.1016/j.asoc.2023.110586>
15. Idrees, S.M.; Eid, M.S.; Awwad, R. Impact of Altering the Bid Selection Method to Below-Average Method: An Agent-Based Modeling Approach. *Journal of Management in Engineering* 2023, 39, 04023003.
<https://doi.org/10.1061/JMENEA.MEENG-5084>
16. de Sousa, L.J.; Poças Martins, J.; Sanhudo, L. Tackling the Data Sourcing Problem in Construction Procurement Using File-Scraping Algorithms. *Engineering Proceedings* 2023, 53, 34.
<https://doi.org/10.3390/IOCBD2023-15190>
17. Egwim, C.N.; Alaka, H.; Demir, E.; et al. Artificial Intelligence in the Construction Industry: A Systematic Review of the Entire Construction Value Chain Lifecycle. *Energies* 2024, 17, 182.
<https://doi.org/10.3390/en17010182>
18. Baghalzadeh Shishehgharkhaneh, M.; Moehler, R.C.; Fang, Y.; Aboutorab, H.; Hijazi, A.A. Construction supply chain risk management. *Automation in Construction* 2024, 162, 105396.
<https://doi.org/10.1016/j.autcon.2024.105396>
19. Abdulla, H.; Baryannis, G. A hybrid multi-criteria decision-making and machine learning approach for explainable supplier selection. *Supply Chain Analytics* 2024, 7, 100074.
<https://doi.org/10.1016/j.sca.2024.100074>
20. Heo, C.; Park, M.; Ahn, C.R. Potential AI-Driven Algorithmic Collusion and Influential Factors in Construction Bidding. *Journal of Computing in Civil Engineering* 2024, 38, 04024016.
<https://doi.org/10.1061/JCCEE5.CPENG-5683>
21. Dzhupova, Z.; Bosch-Mauchand, M.; Doumeingts, G.; Ballot, E. Choosing the right path for AI integration in engineering companies: A strategic approach within the EPC sector. *The Journal of Systems and Software* 2024, 209, 111945.
<https://doi.org/10.1016/j.jss.2023.111945>
22. Liang, D.; Cao, W.; Zhang, Y.; Xu, Z. A two-stage classification approach for AI technical service supplier selection considering stakeholders' bounded rationality. *Information Sciences* 2024, 653, 119762.
<https://doi.org/10.1016/j.ins.2023.119762>

23. Durgam, H.K.R.; Ramani, P.V. Green construction supply chain: Reviewing advancements, obstacles, and possibilities. *Cleaner Logistics and Supply Chain* 2025, 17, 100259. <https://doi.org/10.1016/j.clscn.2025.100259>
24. Liu, C.-C.; Chou, Y.-C. Automated legal consulting in construction procurement using metaheuristically optimized large language models. *Automation in Construction* 2025, 170, 105891. <https://doi.org/10.1016/j.autcon.2024.105891>
25. Dikmen, S.U.; Dikmen, I.; Budayan, C. Automated construction contract analysis for risk and responsibility assessment using natural language processing and machine learning. *Computers in Industry* 2025, 170, 104251. <https://doi.org/10.1016/j.compind.2025.104251>
26. Heo, C.; Ahn, C.R.; Oh, M.; Park, M. Algorithmic Collusion in Construction Bidding: Risks, Dynamics, and Regulatory Implications. *Journal of Management in Engineering* 2025, 41, 04025045. <https://doi.org/10.1061/JMENEA.MEENG-6704>
27. Oo, B.L.; Lim, B.T.H.; Feng, Z.; et al. Predicting the number of bidders in construction competitive bidding using explainable machine learning models. *Construction Innovation* 2025. <https://doi.org/10.1108/CI-10-2024-0325>
28. Serugga, J. AI-Assisted Game Theory Approaches to Bid Pricing Under Uncertainty in Construction. *AppliedMath* 2025, 5, 39. <https://doi.org/10.3390/appliedmath5020039>
29. Hayati, E.N.; Jauhari, W.A.; Damayanti, R.W.; Rosyidi, C.N. An integrated analytics model for supplier selection and order allocation with machine learning and multi-criteria optimization. *Decision Analytics Journal* 2025, 16, 100599. <https://doi.org/10.1016/j.dajour.2025.100599>
30. Ghayur, M.K. Impact of ERP Implementation, Process Automation, and Supplier Governance on Project Cost Efficiency: Evidence from Vision 2030 Projects in Saudi Arabia. *Journal of Humanities and Social Sciences Studies* 2025, 3, 66-83. <https://doi.org/10.61503/JHHSS/v3i2.77>