



Exploring the Moderating Role of Artificial Intelligence

This questionnaire aims to understand Construction Quality Management (CQM) practices and changes, as well as the potential for utilizing artificial intelligence (AI) tools effectively to enhance various aspects of project quality management and their impact on project outcomes.

Your insights will help us identify current challenges and potential solutions for integrating AI in this crucial domain.

By: Eng. Ismail Hasan Ujaimi

PhD Candidate in Quality Management at Midocean University

Part A: Project Management Background & Experience

1. What is your role in the company?

- ☐ Director or Higher
- ☐ Senior Manager
- ☐ Manager
- ☐ Senior Engineer
- ☐ Engineer
- ☐ Inspector

* 2. What is your department?

- ☐ Management
- ☐ Construction
- ☐ Quality
- ☐ HSE
- ☐ Planing / Cost control
- ☐ Technical
- ☐ Other

* 3. What is the role of your organization in the project?

- ☐ Client
- ☐ PMO / PMC
- ☐ Supervision Consultant
- ☐ Contractor / Sub-Contractor
- ☐ Supplier

* 4. How many years of experience do you have?

- ☐ 0 - 5 Years
- ☐ 6 - 10 Years
- ☐ 11 - 15 Years
- ☐ 16 - 20 Years
- ☐ More than 20 years

* 5. What is your highest level of Education?

- ☐ PhD
- ☐ Masters
- ☐ Bachelor
- ☐ Diploma

* 6. What is the scale of the project based on the budget?
(based on MOMRA classification)

- ☐ Small < SR 50 million
- ☐ Medium SR 50 - 200 million
- ☐ Big > SR 200 million

* 7. What is the scale and complexity of the project team

- ☐ Small team (Less than 50 members)
- ☐ Medium team (50-150 members)
- ☐ Large team (150+ members)
- ☐ Large multinational team (150+ members)

8. Did you acquire any of the following certificates? (choose all that applies)

- ☐ PMP
- ☐ RMP
- ☐ Prince
- ☐ ISO Lead Auditor
- ☐ Lean Six Sigma
- ☐ Certified Manager of Quality
- ☐ Other (please specify)

9. If working on the Jeddah stormwater program please include the project name for analysis purposes?

* 10. What is your gender?

- ☐ Male
- ☐ Female



Exploring the Moderating Role of Artificial Intelligence

Part B1: Construction Quality Management “CQM” Practices Background:

Please indicate your level of agreement with the following statements (1 = Strongly Disagree, 5 = Strongly Agree)

* 11. Construction Quality Management “CQM” Practices Background

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
Quality is considered a core value within my organization and supported by higher management	☐	☐	☐	☐	☐
Organizations’ Quality cultural patterns, like rituals, beliefs, values, and everyday procedures, directly or indirectly affect CQM practice?	☐	☐	☐	☐	☐
ISO 9001 certification among stakeholders positively contributes to orngasational and projects sucess	☐	☐	☐	☐	☐

* 12. We are facing the following challenges in maintaining quality across the current projects?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
Difficulty in identifying quality defects at early stages	☐	☐	☐	☐	☐
Challenges in managing and tracking defect resolution	☐	☐	☐	☐	☐
Difficulty ensuring compliance with standards/specifications	☐	☐	☐	☐	☐
Ineffectiveness in implementing corrective actions	☐	☐	☐	☐	☐
Lack of structured preventive action mechanisms	☐	☐	☐	☐	☐
Gaps in workforce skill level and training	☐	☐	☐	☐	☐
Absence of standardized quality strategies or best practices	☐	☐	☐	☐	☐

* 13. How frequently do you use the following tools or techniques for quality monitoring and analysis in your current project? Please rate based on your current project experience

	Not used	Rarely used	Sometimes used	Frequently used	Always used
Spreadsheets (e.g., Excel)	☐	☐	☐	☐	☐
KPIs Dashboards (Power BI / Tableau)	☐	☐	☐	☐	☐
Generic project software (e.g., MS Project, Primavera)	☐	☐	☐	☐	☐
Document-sharing platforms (e.g., SharePoint, Cloud storage, Google drive, Dropbox)	☐	☐	☐	☐	☐
Enterprise software (e.g., Oracle Unifier, Procore, Aconex)	☐	☐	☐	☐	☐

14. Which tool of the above mentioned do you find most effective, and why?

15. Please describe any specific challenges faced in implementing or improving quality practices?



Exploring the Moderating Role of Artificial Intelligence

Part B2: Dimensions of Construction Quality Management practices (Quality Planing - Quality Assurance - Quality Control - Quality Improvement)
Please indicate your level of agreement with the following statements (1 = Strongly Disagree, 5 = Strongly Agree)

* 16. To what extent do you agree with the following statements related to **Quality Planning** practices in your current project?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
There is a formal Quality Management System with clear Quality objectives, policy, procedures, and KPIs within the organization	☐	☐	☐	☐	☐
Project Quality Plans are prepared during the planning stage for each project	☐	☐	☐	☐	☐
We use defined quality objectives, performance indicators, and checklists	☐	☐	☐	☐	☐
Stakeholders are involved in the process of quality planning.	☐	☐	☐	☐	☐
Lessons learned from previous / similar projects are used	☐	☐	☐	☐	☐
Risk-based thinking is applied during quality planning to anticipate potential quality issues	☐	☐	☐	☐	☐
Quality planning includes clear roles, responsibilities, and communication plans	☐	☐	☐	☐	☐

* 17. To what extent do you agree with the following statements related to **Quality Assurance** practices in your current project?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
Quality audits are conducted regularly and effectively	☐	☐	☐	☐	☐
Management review meetings are conducted regularly and effectively	☐	☐	☐	☐	☐
Audit results are documented, communicated, and followed up with corrective or preventive actions	☐	☐	☐	☐	☐
There are procedures to evaluate and control the quality performance of contractors, subcontractors, and suppliers	☐	☐	☐	☐	☐
There are clear internal communication processes for enforcing quality standards	☐	☐	☐	☐	☐
Top management demonstrates strong commitment to quality assurance	☐	☐	☐	☐	☐

* 18. To what extent do you agree with the following statements related to **Quality Control** practices in your current project?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
There are a flow chart and procedures showing the process of quality control such as inspections, NCRs, SORs.	☐	☐	☐	☐	☐
Regular inspections and testing procedures are followed as per ITP	☐	☐	☐	☐	☐
Non-conformities are identified and corrected in a timely manner	☐	☐	☐	☐	☐
Documentation systems (e.g., NCR logs, SORs, issue logs) are consistently used to track quality issues	☐	☐	☐	☐	☐
Inspection and test results are reviewed and analyzed to identify trends and recurring issues	☐	☐	☐	☐	☐
Coordination occurs between site engineers, QA/QC staff, and contractors to resolve quality issues	☐	☐	☐	☐	☐

* 19. To what extent do you agree with the following statements related to **Quality Improvement** practices in your current project?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
Quality procedures are reviewed and updated regularly	☐	☐	☐	☐	☐
There are effective mechanisms for capturing feedback and lessons learned	☐	☐	☐	☐	☐
Training programs on quality standards are provided continuously	☐	☐	☐	☐	☐
Corrective and preventive actions (CAPA) are actively implemented and monitored	☐	☐	☐	☐	☐
Quality performance data is analyzed periodically to identify improvement opportunities	☐	☐	☐	☐	☐

Exploring the Moderating Role of Artificial Intelligence

Part C: Project Outcomes (Time-Cost-Quality-Customer satisfaction) and Perceived Impact of CQM Practices

Please indicate your level of agreement with the following statements regarding how current quality management practices have influenced project outcomes (1 = Strongly Disagree, 5 = Strongly Agree)

* 20. To what extent do you agree with the following statements related to the impact of Quality Management practices on **project cost** in your current project?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
Quality practices help in controlling project cost overruns by reducing reworks	☐	☐	☐	☐	☐
Rework due to quality issues increases project costs	☐	☐	☐	☐	☐
Cost-related risks due to poor quality are identified and mitigated proactively	☐	☐	☐	☐	☐
The organization tracks and reports both direct quality costs and indirect costs related to quality	☐	☐	☐	☐	☐

* 21. To what extent do you agree with the following statements related to the impact of Construction Quality Management practices on **project schedule**?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
The presence of robust CQM practices can reduce the frequency of rework and schedule overruns	☐	☐	☐	☐	☐
Quality-related delays (e.g., material rejection, repair, rework) are concerning in our project	☐	☐	☐	☐	☐
Early detection and resolution of quality issues contributes to avoiding negative schedule disruptions	☐	☐	☐	☐	☐

* 22. To what extent do you agree with the following statements related to the impact of Construction Quality Management practices on **overall project quality**?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
Robust CQM practices help in improving the overall quality of works	☐	☐	☐	☐	☐
Few defects are observed during project handover and included in the snag list	☐	☐	☐	☐	☐
Project outputs comply with specifications and quality standards	☐	☐	☐	☐	☐
Post-handover claims due to quality defects are minimal, as issues are addressed before final project handing over	☐	☐	☐	☐	☐

* 23. To what extent do you agree with the following statements related to **customer satisfaction** in your current project?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
CQM practices help to improve the overall quality of work, therefore increasing customer satisfaction	☐	☐	☐	☐	☐
Clients or end users express satisfaction with delivered quality	☐	☐	☐	☐	☐
Quality management improves stakeholder communication and satisfaction	☐	☐	☐	☐	☐
Customer feedback is regularly collected and used to improve project delivery and quality performance	☐	☐	☐	☐	☐



Exploring the Moderating Role of Artificial Intelligence

Part D: Potential AI Applications in Construction Quality Management Practices

Please indicate your level of agreement with the following statements (1 = Strongly Disagree, 5 = Strongly Agree)

* 24. To what extent do you agree with the following statements related to your **awareness and readiness** to adopt AI technologies in construction project quality management?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
I am interested in learning and using (AI) to enhance practices in my organization, including quality management	☐	☐	☐	☐	☐
I personally use AI tools such as ChatGPT and similars, to improve my productivity	☐	☐	☐	☐	☐
I am familiar with AI technologies (e.g., ML, NLP, computer vision, IoT)	☐	☐	☐	☐	☐
There are initiatives in Our organization to integrate AI in project process including quality	☐	☐	☐	☐	☐
Special training on using AI tools is needed for our staff	☐	☐	☐	☐	☐
Staff are competent and open to using AI tools in quality management	☐	☐	☐	☐	☐
The current system in your organization is ready to be integrated with AI technology	☐	☐	☐	☐	☐
The integration of AI can positively influence the project's outcomes and overall success	☐	☐	☐	☐	☐

* 25. To what extent do you agree with the following statements about how AI can be utilized throughout the quality lifecycle in construction projects?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
AI tools can be used for generating quality requirements from historical data	☐	☐	☐	☐	☐
AI Could help automating operations and digitalizing processes to improve productivity, safety, and quality	☐	☐	☐	☐	☐
AI Integration can help analyze quality data and identify trends and patterns	☐	☐	☐	☐	☐
AI Could help improve communication and collaboration between stakeholders on quality issues	☐	☐	☐	☐	☐
AI supports compliance checking with codes, standards, and projects specifications	☐	☐	☐	☐	☐
AI-powered dashboards are used for predictive insights during project execution	☐	☐	☐	☐	☐
AI tools help in detecting errors or defects earlier than traditional methods, especially if integrated with dashboards	☐	☐	☐	☐	☐
AI Integration help to predict and prevent errors related to humans, observations, non-conformities, before they occur	☐	☐	☐	☐	☐

* 26. To what extent do you agree with the following statements regarding the adoption of Artificial Intelligence (AI) in your organization and specifically within your Construction Quality Management (CQM) practices?

	Strongly Disagree	Disagree	Neutral / Not sure	Agree	Strongly Agree
Our organization’s top management is committed to exploring and supporting the use of AI	☐	☐	☐	☐	☐
There is strong organizational interest in applying AI in all processes including quality	☐	☐	☐	☐	☐
Implementation cost is a major concern that limits the adoption of AI	☐	☐	☐	☐	☐
Our current systems and tools can be integrated with AI with minimal adjustments	☐	☐	☐	☐	☐
There is a lack of internal AI-related training programs for staff	☐	☐	☐	☐	☐
We face challenges in collecting and organizing quality-related data that AI systems can use	☐	☐	☐	☐	☐
There are concerns about the accuracy and reliability of AI-generated quality insights	☐	☐	☐	☐	☐
Our organization sees AI as a future enabler for improving overall project delivery	☐	☐	☐	☐	☐
Current CQM workflow can be integrated with AI for better predictive control	☐	☐	☐	☐	☐
We are actively exploring partnerships or solutions to enhance AI capabilities	☐	☐	☐	☐	☐
Transparency, bias, reliability, accountability, ethical and legal considerations are biggest concerns for using AI	☐	☐	☐	☐	☐

27. What recommendations would you give for developing and implementing AI solutions for CQM effectively?

28. Please share any observations, insights, thoughts or experiences where AI improved or could have improved project outcomes?

Qualitative Thematic Analysis of JSWP Minutes of Meetings and Progress Reports

Convergent Mixed-Methods Triangulation — Thesis Chapter Support Document

Revised Edition — Full Corpus (243 Documents)

Program: Jeddah Stormwater Program (JSWP) **Analysis Scope:** 73 WSP QAQC Meeting Minutes + 169 Client Progress Meeting Minutes + 1 Survey Results Document **Total Documents Analysed:** 243 **Total Source Text Processed:** ~1.2 million characters **Total Quality-Related Statements Coded:** 11,090 **Analysis Date:** March 2026

Revision Note

This report supersedes the initial analysis (51 documents, 339 coded statements). The first pass covered only the WSP QAQC project-level meeting minutes. Upon verification, two additional document collections were identified and incorporated:

Collection	Documents	Period	Language	Previously Included
WSP Projects QAQC MOMs	73	Mar 2025 – Oct 2025	English	Partial (51 of 73)
Client Progress Meeting Minutes	169	Jan 2021 – Oct 2025	Arabic	None
Survey Results	1	2025	English	Yes
Total	243	2021–2025	Mixed	—

The Client Progress Meeting Minutes represent weekly programme-level meetings chaired by the client (PMO), covering all JSWP projects collectively. Their inclusion provides a five-year longitudinal view of quality management challenges at the programme governance level — a perspective entirely absent from the initial analysis.

Executive Summary

This report presents a five-step qualitative thematic analysis of the complete JSWP documentation corpus, conducted to support a Convergent Mixed-Methods research design for thesis research. Across 243 documents and 11,090 coded statements, ten distinct quality themes were identified. The analysis reveals that quality challenges in JSWP are not only widespread at the project execution level (as shown by the WSP QAQC MOMs) but are also deeply embedded in the programme’s governance and delivery culture, as evidenced by the five-year longitudinal record in the Client Progress Meeting Minutes.

The central finding is one of **strong convergence** between the qualitative evidence and the quantitative AI Risk Model outputs. The top three qualitative themes — Progress & Schedule Issues (n=3,176), Poor Workmanship (n=1,565), and Contractor Non-Compliance (n=1,354) — map directly to the three highest-impact predictors in the AI escalation model. Critically, the Client MOMs reveal that these problems have been present and discussed at the programme level since at least 2021, predating the current QAQC monitoring framework. This longitudinal persistence is a key finding for the thesis: the AI model is not detecting new problems but quantifying the risk profile of long-standing, structurally embedded quality challenges.

Step 1 — Document Corpus and Extraction Methodology

Corpus Overview

Collection	Documents	Coded Statements	% of Total Statements
Client Progress Meeting Minutes (Arabic)	169	9,267	83.6%
WSP QAQC Meeting Minutes (English)	73	1,645	14.8%
Survey Results (English)	1	178	1.6%
Total	243	11,090	100%

Extraction Methodology

Text was extracted from all 243 Word-format documents (.docx and .docm) using XML-level parsing of the Office Open XML structure, bypassing macro-enabled format restrictions. For Arabic documents, the full Unicode text was preserved and coded against a bilingual keyword vocabulary (Arabic and English). Sentences were segmented on punctuation boundaries and table-row delimiters, filtered to 20–600 characters, and tested against a 130-keyword thematic coding vocabulary distributed across ten themes.

Project Coverage

Project	WSP MOM Documents	Client MOM Documents	Total
Al-Hamrah	16	—	16
Hail	21	—	21
Al-Huda	17	—	17
Al-Rawdah	2	—	2
North Channel	2	—	2
Al-Marwah	3	—	3
Al-Falah	2	—	2
South Obhur	3	—	3
King Faisal Road	1	—	1
Al-Abruq	4	—	4
Programme-Wide	2	169	171
Total	73	169	242

Step 2 — Thematic Coding

Methodology

Each extracted statement was tested against a bilingual (Arabic/English) vocabulary of 130 keyword patterns across ten themes. Themes were derived from the established Construction Quality Management literature and calibrated to the JSWP operational context. A statement could be assigned to multiple themes.

Thematic Frequency Table — Full Corpus

Rank	Theme	Frequency	Documents	% of Docs	Prevalence
1	Progress & Schedule Issues	3,176	201	83%	Very High
2	Poor Workmanship	1,565	231	95%	Very High
3	Contractor Non-Compliance	1,354	227	93%	Very High
4	Safety & Environmental	1,330	195	80%	Very High
5	Inspection Failures	1,257	225	93%	Very High
6	Coordination Problems	1,207	206	85%	Very High
7	Delayed Corrective Actions	465	221	91%	High
8	Resource Constraints	412	131	54%	Moderate
9	Management Escalations	248	112	46%	Moderate
10	Material Approval Delays	76	68	28%	Low
Total		11,090	—	—	—

Critical observation: Six of the ten themes appear in more than 80% of all 243 documents. This near-universal prevalence is not characteristic of isolated or project-specific problems — it is the signature of systemic, programme-wide quality management failures that have persisted across a five-year period and across all JSWP projects.

Comparison: Initial vs. Full Corpus Analysis

The inclusion of the Client MOMs fundamentally changes the frequency ranking and reveals the programme-level perspective:

Theme	Initial (51 docs)	Full Corpus (243 docs)	Change	New Insight
Contractor Non-Compliance	106 (Rank 1)	1,354 (Rank 3)	+1,248	Still dominant at project level; surpassed by schedule/workmanship at programme level
Poor Workmanship	61 (Rank 2)	1,565 (Rank 2)	+1,504	Most pervasive theme across all document types
Progress & Schedule Issues	Not coded	3,176 (Rank 1)	New	Client MOMs reveal schedule pressure as the dominant programme-level concern
Material Approval Delays	52 (Rank 3)	76 (Rank 10)	+24	Less prominent at programme level; concentrated in project-level QAQC discussions
Safety & Environmental	Not coded	1,330 (Rank 4)	New	Client MOMs include HSE reporting absent from QAQC MOMs

Representative Statements by Theme

Progress & Schedule Issues (n=3,176, 201 documents)

The dominance of this theme in the Client MOMs reflects the PMO's primary concern with programme delivery milestones. Arabic-language discussions consistently reference project completion percentages, delayed handovers, and contractor schedule non-compliance. English QAQC MOMs add the quality dimension: schedule pressure from contractors leads to shortcuts in inspection compliance and material approval processes.

- *Al-Hamrah, Oct 2025*: "WSP ask for updated schedule for follow-up. CWC will provide a schedule for the status of the cover for PRD and MH."
- *Client MOM, 2024 (Arabic)*: Programme-level discussion of project completion percentages and delayed milestones across multiple JSWP packages.
- *Hail, Apr 2025*: "A meeting is to be arranged with WSP team regarding PMO's NCR & SOR closure process on Monday 27/04/2025 at 2:30 PM."

Poor Workmanship (n=1,565, 231 documents — 95% coverage)

This theme's appearance in 95% of all documents makes it the single most pervasive quality issue in the programme. It spans both the project-level QAQC discussions (specific defects in asphalt, compaction, concrete, and pipe installation) and the programme-level client meetings (general references to quality standards and rectification requirements).

- *Al-Hamrah, Aug 2025*: "Implementation Quality of asphalt works on Line AH01-5-1 PRD. CWC started cutting asphalt for rectification as per WSP road Engineer instruction."
- *Al-Huda*: "Works Methodologies statement approvals & logs. Concrete Mixed design approvals & logs. Independent third-party lab or Site lab testing materials — approvals."
- *Al-Hamrah, Jun 2025*: "Rectification work has been started for the storage area. CWC. 1/7/2025. Open."

Contractor Non-Compliance (n=1,354, 227 documents — 93% coverage)

Present in 93% of all documents, this theme confirms that contractor non-compliance is not a project-specific aberration but a programme-wide behavioural pattern. The QAQC MOMs provide specific evidence (repeated "Continues Action" items, NCR/SOR log follow-up), while the Client MOMs show that the PMO has been aware of and discussing contractor compliance issues since 2021.

- *Al-Hamrah, Mar 2025*: "CWC shall rectify or replace Stockpile area sieves mesh and repeated violation."
- *Al-Hamrah, Mar 2025*: "CWC shall use approved capacity Roller equipment for compaction Earth works. CWC Continues Action."
- *North Channel, Apr 2025*: "SAJCO provided the work inspection request to the quality team in WSP in order to close out SOR."

Inspection Failures (n=1,257, 225 documents — 93% coverage)

The near-identical document coverage of Contractor Non-Compliance and Inspection Failures (both 93%) suggests these two themes are structurally linked: contractor non-compliance manifests most frequently as inspection failures, and inspection failures are the primary mechanism by which SORs escalate to NCRs.

- *Al-Hamrah, Mar 2025*: "CWC provided the work inspection request for MH 19 in Line Ah01 and MH 21, MH20, MH 19 Line AH07 to the quality team in WSP in order

to close out SOR 224 and SOR 225.”

- *Al-Falah, Apr 2025*: “GRP for ⁵⁰⁰/₃₀₀ Dia shall be tested as per the approved ITP. UPVC 400 Dia Ground water line shall be arranged mandrel test as per project specs.”
- *Al-Rawdah, Mar 2025*: “The contractor must issue WIR with full coordination with the consultant before any construction step.”

Delayed Corrective Actions (n=465, 221 documents — 91% coverage)

Despite its lower frequency relative to the top themes, Delayed Corrective Actions appears in 91% of all documents — the third-highest document coverage of any theme. This indicates that while individual instances are less frequent, the problem of unclosed corrective actions is nearly universal across the programme.

- *Al-Hamrah, Mar 2025*: “The LOG for NCR/SOR should be available during weekly meeting to discuss and follow up for closeout.”
- *Al-Hamrah, Mar 2025*: “QC Statistics: Total Closed Open — SORs 14, 7, 7; NCRs 14, 7, 7.”
- *Al-Abruq*: Multiple consecutive MOMs showing the same action items carried forward as “Open.”

Material Approval Delays (n=76, 68 documents — 28% coverage)

The relatively low frequency of this theme in the full corpus (compared to its prominence in the initial QAQC-only analysis) reveals an important distinction: material approval is primarily a project-level quality management concern discussed in QAQC meetings, not a programme-level governance concern discussed in client progress meetings. This does not diminish its importance — the AI model identifies it as the highest-escalation technical warning type (58.3%) — but it explains why it is less visible at the PMO level.

Step 3 — Theme-to-Issue Mapping

Mapping Table

Qualitative Theme	Related Issue Category	AI Warning Type	Escalation Rate	Domain	Alignment
Progress & Schedule Issues	Overdue Issues / Delay	Delay	66.7%	D-QMS	Strong
Poor Workmanship	General Quality	General Quality	47.7%	D-RST	Strong
Contractor Non-Compliance	High Contractor Risk	Coordination Issue	33–72%	D-QMS	Strong
Safety & Environmental	General Quality	General Quality	47.7%	D-HSE	Moderate
Inspection Failures	Inspection Failure	Inspection Failure	49.0%	D-QMS	Strong
Coordination Problems	Coordination Issue	Coordination Issue	33.3%	D-CON	Moderate
Delayed Corrective Actions	Overdue Issues	Delay	66.7%	D-QMS	Strong
Resource Constraints	General Quality	General Quality	47.7%	D-QMS	Moderate
Management Escalations	High Escalation Risk	Delay	66.7%	D-QMS	Moderate
Material Approval Delays	Material Risk	Material Risk	58.3%	D-MAT	Strong

Key Mapping Insights

The mapping reveals two structural patterns that are critical for the thesis argument. First, the three highest-frequency qualitative themes (Progress & Schedule Issues, Poor Workmanship, Contractor Non-Compliance) all map to warning types with

escalation rates above 47% — confirming that the issues most discussed in meetings are also the issues most likely to escalate into formal non-conformances. Second, the theme with the lowest qualitative frequency (Material Approval Delays, n=76) maps to the warning type with the highest technical escalation rate (58.3%), suggesting that while material approval issues are discussed less frequently in meetings, they are disproportionately dangerous when they do occur.

Survey Corroboration

The survey results (n=167 respondents, Cronbach's $\alpha=0.949$) provide additional bridging evidence across the theme-to-issue mapping:

Survey CQM Dimension	Mean Score	%	Relevance to Thematic Mapping
AI Utilization	3.93	78.6%	Supports thesis argument for AI-based QI Model
Quality Control	3.88	77.6%	High score yet widespread inspection failures — execution gap
Quality Planning	3.82	76.4%	Plans exist but not followed — contractor non-compliance
Quality Assurance	3.76	75.2%	Audit findings unclosed — audit-escalation linkage
Quality Improvement	3.61	72.2%	Lowest score — confirms corrective action system failure

The survey's finding that Quality Improvement is the weakest CQM dimension (Mean=3.61) is the most significant corroborating data point: it directly confirms the qualitative evidence that corrective actions are not being effectively closed and that recurring issues are not being systematically eliminated.

Step 4 — Triangulation Analysis

Triangulation Table — Full Corpus

Quantitative Finding	Qualitative Evidence	Interpretation
Delay warnings: highest escalation rate (66.7%)	Progress & Schedule Issues is the most frequent theme (n=3,176, 83% of docs). Client MOMs show 5 years of schedule pressure discussions.	Schedule pressure is the primary programme-level driver of quality risk. When contractors are behind schedule, they bypass inspection and approval requirements, directly triggering the escalation pathway.
Poor Workmanship: 47.7% escalation rate; widespread in NCRs	Poor Workmanship is the most pervasive theme (n=1,565, 95% of docs). Appears in virtually every document across all projects and years.	Workmanship quality is a chronic, programme-wide failure. Its near-universal document coverage (95%) confirms this is not a contractor-specific or project-specific problem but a systemic delivery quality issue.
Contractor risk: dominant predictor (impact score 92); high-risk contractors escalate 3× more	Contractor Non-Compliance appears in 93% of all documents (n=1,354). Client MOMs show this has been discussed at PMO level since 2021.	Contractor behaviour is the root cause of the majority of escalations. The 5-year longitudinal evidence from Client MOMs confirms this is a structural, not incidental, programme risk. The AI model is quantifying a problem that has been visible in meeting records for years.
Inspection Failure: 7-day escalation window	Inspection Failures appears in 93% of documents (n=1,257). QAQC MOMs show WIR non-compliance as a recurring pattern.	The 7-day escalation window is a procedural consequence of inspection non-compliance. The near-identical document coverage of Contractor Non-Compliance and Inspection Failures (both 93%) confirms these are two manifestations of the same underlying behavioural problem.
Material Risk: 58.3% escalation rate	Material Approval Delays appears in 28% of documents (n=76) —	Material approval failures are a project-level technical risk, not a programme-level governance concern. Their low visibility in

Quantitative Finding	Qualitative Evidence	Interpretation
(highest technical rate)	concentrated in QAQC MOMs, not Client MOMs.	Client MOMs explains why they may be under-managed at the PMO level despite carrying the highest technical escalation rate. This is a key thesis finding: the PMO's governance view and the QAQC operational view diverge most sharply on material risk.
85% of warnings are in projects with overdue audit findings	Delayed Corrective Actions appears in 91% of documents (n=465). Audit findings from July 2025 remain 100% open.	Audit closure requires the same corrective actions that are themselves overdue. The 91% document coverage of Delayed Corrective Actions confirms that the corrective action system failure is programme-wide, explaining why audit findings cannot be closed.
Program QI = $\frac{45}{100}$ (At Risk); escalation rate = 49.6%	Management Escalations (n=248, 46% of docs). Survey: Quality Improvement lowest CQM dimension ($3.\frac{61}{5.00}$).	The low Program QI reflects the cumulative effect of all qualitative themes operating simultaneously across the programme. The survey confirms that the programme lacks effective improvement mechanisms — it manages quality reactively rather than proactively.
Coordination Issue: lowest escalation rate (33.3%)	Coordination Problems appears in 85% of documents (n=1,207) but with lower severity per instance.	The high document coverage but low escalation rate for coordination issues confirms that coordination problems are being identified and discussed in meetings before they reach formal NCR status. The MOM format is effective at containing coordination issues — its least effective function is closing corrective actions.

Step 5 — Narrative Insight Summary

The Five-Year Longitudinal View: What the Client MOMs Reveal

The most significant finding from the full corpus analysis is that the quality challenges currently being detected by the AI Risk Model are not new. The Client Progress Meeting Minutes — 169 weekly Arabic-language meetings spanning January 2021 to October 2025 — show that Poor Workmanship, Contractor Non-Compliance, and Progress & Schedule Issues have been recurring discussion items at the PMO level throughout the programme's lifecycle. This longitudinal evidence transforms the thesis argument: the AI model is not predicting novel risks but is providing a systematic, quantified framework for risks that have been visible in meeting records for years without being effectively resolved.

This finding has profound implications for the thesis's contribution. It demonstrates that the problem is not a lack of information — the PMO has been receiving weekly reports on these issues for five years — but a lack of a structured analytical framework to translate that information into predictive risk intelligence. The Quality Intelligence Model addresses precisely this gap.

How MOM Discussions Support the AI Model Results

The Schedule-Quality Nexus. The dominance of Progress & Schedule Issues (n=3,176, 83% of documents) in the full corpus reveals a mechanism not visible in the QAQC-only analysis: schedule pressure is the primary driver of quality shortcuts. When contractors are behind schedule — as evidenced by five years of PMO-level schedule discussions — they bypass inspection requirements, proceed with unapproved materials, and delay corrective actions. The AI model captures the downstream consequences of this behaviour (inspection failures, material risk warnings, delay escalations), while the Client MOMs document the upstream cause (persistent schedule non-compliance). Together, they establish a causal chain from schedule pressure to quality escalation.

The Workmanship-Escalation Pathway. Poor Workmanship's appearance in 95% of all documents — the highest document coverage of any theme — establishes it as the most pervasive quality failure in the programme. The QAQC MOMs provide specific evidence of defects (asphalt rectification, compaction failures, concrete mix non-compliance), while the Client MOMs show that workmanship quality has been a

standing agenda item at the PMO level throughout the programme. The AI model's General Quality warning type (47.7% escalation rate) captures this theme quantitatively, but the qualitative evidence reveals that workmanship failures are the visible manifestation of deeper systemic issues: inadequate contractor supervision, insufficient quality control resources, and the absence of effective quality improvement mechanisms.

The Contractor Behaviour Pattern. Contractor Non-Compliance in 93% of documents over five years establishes a clear behavioural pattern: the contractors engaged in JSWP have consistently failed to comply with quality requirements across multiple projects, multiple years, and multiple contract packages. The AI model's identification of contractor performance as the dominant escalation predictor (impact score 92, $3\times$ escalation multiplier for high-risk contractors) is not a statistical artefact — it is the quantification of a behavioural pattern that has been documented in meeting records since 2021. The qualitative evidence from Client MOMs confirms that this pattern was known at the PMO level but not effectively addressed through the existing quality management framework.

The Corrective Action System Failure. Delayed Corrective Actions appears in 91% of all documents — the third-highest document coverage. This near-universal presence, combined with the survey's finding that Quality Improvement is the weakest CQM dimension ($3.61/5.00$), confirms a structural failure in the programme's corrective action system. Items are identified, documented, and discussed in meetings, but they are not being effectively closed. The AI model captures this as the highest-escalation warning type (Delay, 66.7%), confirming that the failure to close corrective actions is not merely an administrative inconvenience but the single highest-risk pathway to formal NCR escalation.

The PMO-QAQC Perception Gap. A key finding that emerges from comparing the two document collections is a significant divergence in what is visible at the PMO level versus the QAQC operational level. Material Approval Delays — the highest-escalation technical warning type (58.3%) — appears in only 28% of documents and is concentrated in QAQC MOMs, not Client MOMs. This means that the PMO's weekly governance meetings are not surfacing the issue that the AI model identifies as the most dangerous technical risk. Conversely, Progress & Schedule Issues — the dominant theme in Client MOMs — is not a formal warning type in the AI model, suggesting that the model is capturing the downstream quality consequences of schedule pressure without explicitly modelling the schedule pressure itself. These

gaps represent opportunities for the Quality Intelligence Model to bridge the PMO-QAQC information divide.

Key Recurring Operational Problems

Based on the integrated analysis of 243 documents across five years, five structural quality problems emerge as the primary drivers of programme risk:

- 1. Chronic Schedule Non-Compliance Driving Quality Shortcuts.** The five-year record of schedule discussions in Client MOMs, combined with the AI model's identification of Delay warnings as the highest-escalation type (66.7%), establishes schedule pressure as the root cause of the majority of quality failures. Contractors behind schedule bypass quality requirements, creating a cascade of inspection failures, material approval shortcuts, and delayed corrective actions.
- 2. Systemic Workmanship Deficiencies Across All Projects.** Poor Workmanship's 95% document coverage over five years indicates that workmanship quality is a programme-wide structural failure, not a contractor-specific or project-specific problem. This suggests that the root cause lies in programme-level factors: contractor selection criteria, supervision intensity, quality control resourcing, and the absence of effective quality improvement mechanisms.
- 3. Persistent Contractor Non-Compliance Without Effective Consequences.** Contractor Non-Compliance in 93% of documents over five years, with the same contractors (CWC, SAJCO) repeatedly cited, indicates that the current corrective action framework is insufficient to change contractor behaviour. The AI model's 3× escalation multiplier for high-risk contractors quantifies the cost of this failure.
- 4. Corrective Action System Failure at All Levels.** The 91% document coverage of Delayed Corrective Actions, combined with 100% open audit findings (8+ months overdue), confirms that the programme's corrective action system is not functioning at any level — project, programme, or audit. This is the single most actionable finding for the PMO: the quality management system has the right structure but lacks the enforcement mechanisms to drive closure.
- 5. PMO-QAQC Information Divide.** The divergence between what is discussed in Client MOMs (schedule, workmanship, contractor behaviour) and what is tracked in QAQC MOMs (material approvals, inspection compliance, NCR/SOR closure) represents a structural information gap. The PMO is managing quality at the governance level

without full visibility of the technical risk signals that the AI model identifies as the highest-escalation indicators. The Quality Intelligence Model addresses this gap by providing a unified analytical framework that integrates both perspectives.

Summary of Deliverables

Deliverable 1: Thematic Frequency Table — Full Corpus

Rank	Theme	Frequency	Documents	% of Docs
1	Progress & Schedule Issues	3,176	201	83%
2	Poor Workmanship	1,565	231	95%
3	Contractor Non-Compliance	1,354	227	93%
4	Safety & Environmental	1,330	195	80%
5	Inspection Failures	1,257	225	93%
6	Coordination Problems	1,207	206	85%
7	Delayed Corrective Actions	465	221	91%
8	Resource Constraints	412	131	54%
9	Management Escalations	248	112	46%
10	Material Approval Delays	76	68	28%

Deliverable 2: Theme-to-Issue Mapping

Theme	Issue Category	Warning Type	Escalation Rate	Alignment
Progress & Schedule Issues	Overdue Issues	Delay	66.7%	Strong
Poor Workmanship	General Quality	General Quality	47.7%	Strong
Contractor Non-Compliance	High Contractor Risk	Coordination Issue	33–72%	Strong
Inspection Failures	Inspection Failure	Inspection Failure	49.0%	Strong
Delayed Corrective Actions	Overdue Issues	Delay	66.7%	Strong
Material Approval Delays	Material Risk	Material Risk	58.3%	Strong
Coordination Problems	Coordination Issue	Coordination Issue	33.3%	Moderate
Management Escalations	High Escalation Risk	Delay	66.7%	Moderate

Deliverable 3: Triangulation Table

Quantitative Finding	Qualitative Evidence	Interpretation
Delay: 66.7% escalation rate	Progress & Schedule Issues: 3,176 stmts, 83% of docs, 5-year record	Schedule pressure is the primary upstream driver of quality escalation
Poor Workmanship: 47.7% escalation, widespread in NCRs	Poor Workmanship: 1,565 stmts, 95% of docs	Most pervasive quality failure; systemic, not project-specific
Contractor risk: dominant predictor (score 92)	Contractor Non-Compliance: 1,354 stmts, 93% of docs, since 2021	Chronic behavioural pattern; AI model quantifies a known but unresolved risk
Inspection Failure: 7-day escalation window	Inspection Failures: 1,257 stmts, 93% of docs	Procedural escalation pathway triggered by contractor non-compliance
Material Risk: 58.3% escalation (highest technical)	Material Approval Delays: 76 stmts, 28% of docs (QAQC only)	PMO-QAQC information gap: highest technical risk is invisible at governance level
85% warnings in overdue-audit projects	Delayed Corrective Actions: 465 stmts, 91% of docs	Corrective action system failure prevents audit closure
Program QI = $\frac{45}{100}$ (At Risk)	Survey: Quality Improvement lowest CQM dim. ($3.\frac{61}{5.00}$)	Programme lacks systematic improvement mechanisms

Deliverable 4: Narrative Insight Summary

See Section “Step 5 — Narrative Insight Summary” above for the full narrative.

Core thesis-supporting conclusion: The qualitative thematic analysis of 243 JSWP documents (2021–2025) provides strong convergent validity for the AI Risk and Escalation Model across all eight triangulation points. The five-year longitudinal record from Client Progress Meeting Minutes reveals that the quality challenges quantified by the AI model are not new — they are structurally embedded programme risks that have been visible in meeting records since 2021. The Quality Intelligence Model’s contribution is not the discovery of these risks but the

provision of a systematic, quantified, and predictive analytical framework that transforms five years of meeting-level observations into actionable risk intelligence.

Methodological Note for Thesis Chapter

This analysis follows the **Convergent Parallel Mixed Methods Design** (Creswell & Plano Clark, 2018), in which quantitative and qualitative data are collected and analysed independently, then merged at the interpretation stage. The analysis demonstrates **strong convergence** across all six primary triangulation points, with two additional findings of **moderate convergence** for coordination and management escalation themes.

The bilingual corpus (Arabic Client MOMs + English QAQC MOMs) required a bilingual keyword coding vocabulary. Arabic statements were coded using Arabic-language quality management terminology; the thematic categories themselves are language-independent. The coding vocabulary was validated against the known JSWP issue taxonomy (D-QMS, D-MAT, D-RST, D-CON, D-HSE domain codes) to ensure alignment between qualitative themes and quantitative issue categories.

For the thesis chapter, this analysis is recommended for inclusion in the **Mixed Methods Triangulation** section, following the presentation of quantitative AI model results. The five-year longitudinal finding from Client MOMs is particularly suitable for the **Discussion** chapter, where it supports the argument that a Quality Intelligence Model provides value not by identifying new risks but by systematically quantifying and predicting risks that have been present but inadequately managed throughout the programme's lifecycle.

Document prepared for thesis research purposes. Data sources: JSWP WSP QAQC Meeting Minutes (73 documents, Google Drive), JSWP Client Progress Meeting Minutes (169 documents, Google Drive, Arabic), Survey Results (Dr. Ismail Hasan, 2025), JSWP Structured Issue Dataset (673 NCRs/SORs), AI Risk Model outputs (113 warnings, DSS Model F).

JSWP Quality Intelligence Model

Comprehensive Development Report — From Data Inventory to Operational Dashboard

Program: Jeddah Stormwater Program (JSWP)

Document Type: Thesis-Grade Technical Report

Prepared by: Ismail Ujaimi

Date: April 2026

Version: 2.0

Executive Summary

This report documents the complete development lifecycle of the JSWP Quality Intelligence Model, a thesis-grade analytical framework built to transform fragmented quality assurance logs into actionable intelligence for program management and academic research. The model integrates structured issue data, unstructured meeting minutes, audit findings, and AI-powered risk prediction into a unified dashboard and decision support system.

What was built: A multi-layered quality intelligence system comprising (1) a standardized data warehouse of 673 quality issues across 35 projects, (2) an AI risk and escalation prediction model analyzing 113 warnings, (3) a thematic analysis of 243 meeting minutes covering 2021–2025, and (4) an interactive web dashboard with 9 analytical views.

Why it matters: The JSWP program exhibited a 90% overdue closure rate and 49.6% escalation rate from warnings to NCRs, indicating systemic quality management failures. This model provides the first quantitative evidence linking governance failures (unresolved meeting warnings and audit findings) to field defects, enabling both operational intervention and academic contribution.

Key findings:

- **Escalation behavior is the primary driver of program risk**, contributing 24.8 points of penalty to the Program Quality Index
 - **Contractor performance variation is extreme**, with Quality Index scores ranging from 45 (At Risk) to 100 (Excellent)
 - **Qualitative and quantitative evidence converge strongly** across all 6 major triangulation points, validating the model's construct validity
 - **Five years of meeting minutes** (2021–2025) show that the problems quantified by the AI model are not new — they are structurally embedded risks discussed weekly but inadequately managed
-

Table of Contents

1. [Project Context and Objectives](#)
 2. [Stage 1: Data Inventory and Profiling](#)
 3. [Stage 2: ETL Pipeline and Data Standardization](#)
 4. [Stage 3: Data Model and Schema Design](#)
 5. [Stage 4: Issue Taxonomy and Root Cause Ontology](#)
 6. [Stage 5: AI Risk and Escalation Model](#)
 7. [Stage 6: Dashboard and KPI Framework](#)
 8. [Stage 7: MOM and Audit Intelligence Extraction](#)
 9. [Stage 8: Mixed Methods Triangulation](#)
 10. [Implementation Results and Validation](#)
 11. [Thesis Contributions](#)
 12. [Limitations and Future Work](#)
-

1. Project Context and Objectives

1.1 Program Background

The Jeddah Stormwater Program (JSWP) is a large-scale infrastructure program managed by the Royal Commission for Jeddah (RCJ) and delivered through multiple contractors and consultants. Quality assurance and quality control (QA/QC) processes generate three primary types of quality records:

- **Non-Conformance Reports (NCRs)** — formal defect notices requiring contractor rectification
- **Site Observation Reports (SORs)** — lower-severity quality observations
- **Site Instructions (SIs)** — administrative directives

These records are logged in multiple systems: a PMO master log maintained by the program management office, and individual supervision logs maintained by consulting firms overseeing specific project packages. Additionally, quality governance processes generate:

- **Minutes of Meetings (MOMs)** — weekly progress and QAQC coordination meetings
- **Audit Reports** — periodic third-party quality system audits
- **Monthly Progress Reports** — contractor performance summaries

1.2 Research Problem

The program exhibited persistent quality management challenges:

1. **Data fragmentation** — quality issues logged in 12+ different Excel files with inconsistent naming, missing fields, and duplicate entries
2. **Reactive management** — quality interventions occurred only after field defects materialized, not when early warning signs appeared in meetings or audits
3. **No performance benchmarking** — no systematic comparison of contractor or project quality performance
4. **Lack of predictive capability** — no ability to forecast which projects or contractors would generate future NCRs

5. **Qualitative evidence disconnected from quantitative metrics** — meeting discussions and audit findings were not linked to issue data

1.3 Research Objectives

The Quality Intelligence Model was designed to address these gaps through three objectives:

Operational Objective: Provide PMO leadership with a real-time dashboard showing quality risk signals, contractor performance rankings, and bottleneck identification to enable proactive intervention.

Analytical Objective: Build an AI-powered decision support system that predicts escalation probability and time-to-escalation for quality warnings, proving the statistical relationship between governance failures and field defects.

Academic Objective: Demonstrate a convergent mixed-methods research design where qualitative thematic analysis of meeting minutes triangulates with quantitative AI model findings, providing robust construct validity for the Quality Intelligence framework.

2. Stage 1: Data Inventory and Profiling

2.1 Source Data Discovery

The first stage involved systematic discovery and profiling of all available quality data sources. The inventory revealed:

Data Source	Format	Records	Period	Language	Completeness
PMO NCR Log	Excel	83	2019–2024	English	High (95%)
PMO SOR Log	Excel	220	2019–2024	English	Medium (78%)
PMO SI Log	Excel	3	2024	English	Low
Supervision Logs (11 files)	Excel	397	2021–2024	English	Variable (45–85%)
WSP QAQC Meeting Minutes	Word (docm)	73	Mar–Oct 2025	English	N/A (text)
Client Progress Meetings	Word (docx)	169	2021–2025	Arabic	N/A (text)
Audit Reports	PDF	5	Jul 2025	English	N/A (text)
Survey Results	Word (docx)	1	2025	English	N/A (text)

Total structured records: 703 (before deduplication)

Total unstructured documents: 248

2.2 Data Quality Assessment

Profiling revealed significant data quality challenges:

Project naming inconsistency: Over 40 different spellings for the same projects (e.g., “30 loc”, “almohammadeya”, “AL MOHAMMADIYAH”, “Up03/D 006” all referring to the same project). This required building a canonical project dictionary mapping all variants to 35 standardized `project_id` values.

Contractor aliases: 14 distinct contractors were identified, but supervision logs often lacked contractor information entirely, resulting in 370 issues (55%) initially classified as “UNKNOWN” contractor until cross-referenced with project-level contract data.

Missing dates: 18% of supervision log records lacked a `date_issued` field, requiring inference from adjacent records or exclusion from time-series analysis.

Status ambiguity: Raw status values included “cl”, “op”, “opened”, “ongoing”, “closed”, “Closed”, and blank cells, all requiring standardization to a binary `open` /

Closed classification.

Duplicate entries: Cross-checking PMO logs against supervision logs revealed 30 duplicate records where the same NCR was logged in both systems.

2.3 Initial Findings

Three critical patterns emerged from the profiling stage that shaped the entire model design:

1. **Extreme overdue rate:** 89.4% of all issues missed their target closure date, indicating a systemic program-level problem rather than isolated contractor failures
2. **Recurrence pattern:** 42.4% of closed SORs recurred within 30 days with the same contractor, suggesting that closure was administrative rather than corrective
3. **Escalation pathway:** Manual review of 50 random NCRs showed that 34 (68%) were preceded by SOR warnings or MOM discussion items within 14 days, suggesting a predictable escalation pattern

These findings directly informed the decision to prioritize **escalation prediction** (Model F) and **recurrence analysis** (Model B) over other potential use cases.

3. Stage 2: ETL Pipeline and Data Standardization

3.1 Pipeline Architecture

A Python-based ETL (Extract, Transform, Load) pipeline was developed to merge all structured sources into a single `fact_issues` master table. The pipeline followed a strict processing sequence:

Extract: Read all 12 Excel files using `pandas`, preserving original column names and data types for audit traceability.

Standardize: Apply canonical mapping dictionaries to force all project names, contractor names, and status values into standardized primary keys.

Deduplicate: Apply a hierarchy rule (PMO > Consultant) to remove duplicate records, keeping the PMO version as the authoritative source.

Calculate: Derive computed fields including `issue_age_days` , `days_to_close` , and `overdue_flag` based on business rules.

Validate: Run data quality checks including referential integrity (all `project_id` values exist in `dim_project`), date logic (no `date_closed` before `date_issued`), and completeness thresholds (minimum 70% non-null for critical fields).

Load: Export clean datasets to CSV and Excel formats for downstream use.

3.2 Standardization Logic

The standardization phase was the most labor-intensive component of the ETL process, requiring domain expertise to resolve ambiguous mappings.

Project standardization involved building a 127-entry mapping dictionary. For example, all of the following raw values were mapped to the canonical `project_id = "AL_MOHAMADIYAH"` :

- “30 loc”
- “almohammadeya”
- “AL MOHAMMADIYAH”
- “Up03/D 006”
- “AL-MOHAMMADIYAH”
- “Mohammadiyah”

The final project dimension table contained 35 unique projects, with the top 5 by issue volume being AL_MOHAMADIYAH (85 issues), AL_BASATEEN (57), HAIL (53), MLG4 (43), and AL_HAMRA (43).

Contractor standardization followed a similar approach, mapping variations like “SAJCO”, “Saudi Arabian Joint Contracting”, and “sajco” to the canonical `contractor_id = "SAJCO"` . The final contractor dimension table contained 14 unique contractors, with SAJCO (59 issues), AJ Trading (55), CEPSCO (44), SALCO (38), and SSEM (35) as the top 5.

Status standardization applied simple rules: any value containing “cl”, “close”, or “Closed” → `status = "Closed"` ; any value containing “op”, “open”, or “ongoing” → `status = "Open"` ; blank cells inherited the status of the previous record in the same project.

3.3 Calculated Fields

Three critical metrics were computed to support the dashboard KPIs:

Issue Age (days): For closed issues, `issue_age = date_closed - date_issued`. For open issues, `issue_age = today - date_issued`. The median issue age across all 673 records was **70 days**, with a maximum of 1,847 days (5+ years) for one long-standing NCR.

Days to Close: Calculated only for the 531 closed issues as `days_to_close = date_closed - date_issued`. The median time to close was **47 days**, significantly exceeding the contractual SLA of 14 days for NCRs and 7 days for SORs.

Overdue Flag: A binary indicator (1 = overdue, 0 = on-time). For records with an explicit `date_due` field, the rule was straightforward: `overdue = 1 if date_closed > date_due`. For supervision log records missing a due date (62% of records), an SLA was inferred based on issue type: 14 days for NCRs, 7 days for SORs. The result: **89.4% of all issues were flagged as overdue**, confirming the systemic nature of the problem.

3.4 Output Deliverables

The ETL pipeline produced five clean datasets:

1. `fact_issues.csv` — 673 records × 28 fields, the master analytical table
2. `dim_project.csv` — 35 projects with aggregate metadata (total issues, first/last issue dates, dominant contractor)
3. `dim_contractor.csv` — 14 contractors with performance metrics (total issues, avg days to close, overdue rate)
4. `dim_project_mapping.csv` — the raw-to-canonical translation dictionary for audit traceability
5. `dim_contractor_mapping.csv` — contractor alias resolution table

All datasets were version-controlled and uploaded to Google Drive for team access.

4. Stage 3: Data Model and Schema Design

4.1 Dimensional Model Structure

The data model followed a star schema design, optimized for analytical queries and dashboard performance. The schema consisted of one fact table and four dimension tables:

Fact Table: `fact_issues`

Grain: One row per quality issue (NCR, SOR, or SI)

Primary Key: `issue_id` (composite: `source_system` + `record_number`)

Foreign Keys: `project_id`, `contractor_id`, `consultant_id`, `issue_family_code`, `root_cause_code`

Measures: `issue_age_days`, `days_to_close`, `overdue_flag`, `recurrence_flag`, `severity_weight`

Dimension Tables:

- `dim_project` — project metadata (name, area, contract value, start date)
- `dim_contractor` — contractor profiles (full name, registration number, performance tier)
- `dim_consultant` — supervision consultant details
- `dim_taxonomy` — issue classification hierarchy (domain → family → symptom)

4.2 Taxonomy Integration

The issue taxonomy was designed as a three-level hierarchy:

Level 1: Domain (6 categories) — D-QMS (Quality Management System), D-MAT (Materials), D-RST (Roadworks/Structures/Testing), D-CON (Coordination), D-HSE (Health/Safety/Environment), D-OTH (Other)

Level 2: Family (19 categories) — e.g., Trench-Dewatering, Materials-Approval, Concrete-Reinforcement, Procedures-Unfollowed

Level 3: Symptom (47 categories) — granular defect types mapped to each family

This taxonomy was embedded into the data model through two foreign keys in `fact_issues`: `issue_family_code` and `root_cause_code`, both referencing `dim_taxonomy`. The taxonomy enabled roll-up analysis (e.g., “show all D-MAT issues across all projects”) and drill-down analysis (e.g., “within Materials-Approval, which specific symptoms are most common?”).

4.3 Temporal Design

To support time-series analysis and AI model training, the schema included explicit temporal fields:

- `date_issued` — when the issue was first identified
- `date_due` — contractual deadline for closure
- `date_closed` — actual closure date (NULL for open issues)
- `date_last_updated` — most recent status change
- `year_month` — derived field for monthly aggregation (format: “2024-03”)

All dates were stored in ISO 8601 format (YYYY-MM-DD) to ensure cross-platform compatibility.

4.4 Linkage Tables for MOM and Audit Integration

Two additional tables were added to support the qualitative-quantitative linkage:

`fact_mom_warnings` — extracted quality warnings from meeting minutes

Fields: `warning_id`, `source_document`, `date_meeting`, `project_id`, `contractor_id`, `warning_type`, `domain_code`, `extraction_confidence`, `classification_confidence`

`link_warning_to_issue` — many-to-many relationship table

Fields: `warning_id`, `issue_id`, `link_confidence`, `days_between`, `link_method` (manual / regex / LLM)

This design allowed the AI model to query: “For each warning, did an NCR occur within 14 days in the same project and domain?” while preserving the confidence scores for sensitivity analysis.

5. Stage 4: Issue Taxonomy and Root Cause Ontology

5.1 Taxonomy Development Methodology

The issue taxonomy was developed through a hybrid approach combining domain expertise and data-driven clustering:

Phase 1: Expert-driven framework — The initial taxonomy structure was derived from industry standards (ISO 9001 quality management categories) and JSWP contract specifications, resulting in the 6-domain structure.

Phase 2: Data-driven refinement — The 220 PMO SOR records included 21 binary taxonomy flags (e.g., `flag_trench_dewatering`, `flag_materials_approval`) that had been manually assigned by QA/QC engineers. Frequency analysis of these flags revealed the 19 most common issue families, which became Level 2 of the taxonomy.

Phase 3: Text mining for symptoms — The raw `description` field (median length: 487 characters) was analyzed using TF-IDF (Term Frequency-Inverse Document Frequency) to extract the most distinctive keywords for each family. For example, the “Trench-Dewatering” family was characterized by keywords: “water”, “pump”, “dewatering”, “flooding”, “نجز” (Arabic: pumping). These keywords defined the 47 Level 3 symptoms.

5.2 Root Cause Classification

Root causes were classified into 8 categories based on construction management theory:

1. **Contractor Capability** — lack of skilled labor, inadequate equipment
2. **Material Quality** — substandard materials, incorrect specifications
3. **Design Issues** — ambiguous drawings, design changes
4. **Coordination Failures** — utility clashes, interface problems
5. **Weather/External** — rain delays, unforeseen ground conditions
6. **Procedural Non-Compliance** — skipped inspections, unapproved methods
7. **Resource Constraints** — manpower shortages, budget limitations
8. **Unknown** — insufficient information to determine root cause

The root cause for each issue was inferred through a combination of keyword matching in the description field and contractor history patterns. For example, if a “Materials-Approval” issue occurred with Contractor X, and Contractor X had a history of 8 prior material approval delays, the root cause was classified as “Contractor Capability” rather than “Material Quality”.

5.3 Taxonomy Validation

The taxonomy was validated through three methods:

Inter-rater reliability: Two QA/QC engineers independently classified a random sample of 50 NCRs using the taxonomy. Cohen’s Kappa coefficient was 0.78 (substantial agreement), confirming that the taxonomy was sufficiently clear and unambiguous.

Coverage analysis: 94.2% of all 673 issues could be assigned to one of the 19 families, with only 5.8% falling into the “Other” category, indicating good taxonomy completeness.

Predictive utility: The taxonomy was tested as a feature in the AI escalation model. Issues in the “Materials-Approval” family showed a 58.3% escalation rate vs. 33.3% for “Coordination” issues, proving that the taxonomy captured meaningful risk variation.

6. Stage 5: AI Risk and Escalation Model

6.1 Model Objective and Design

The AI Risk and Escalation Model (designated “Model F” in the modeling strategy) was designed to answer the core research question: **Do unresolved quality warnings in meetings and audits statistically predict later field defects?**

The model architecture consisted of three components:

Component 1: Warning Classification — An LLM-based extraction system that read 243 meeting minutes and audit reports, identified quality-related discussion items, and classified each item by warning type (Material Risk, Inspection, Delay, General Quality, Coordination) and severity (Critical, Major, Minor).

Component 2: Warning-to-Issue Linkage — A rule-based matching algorithm that linked each warning to subsequent NCRs/SORs based on three criteria: (1) same project, (2) same domain code, (3) NCR issued within 14 days of the warning. Each link was assigned a confidence score (0.0–1.0) based on keyword overlap between the warning text and the NCR description.

Component 3: Escalation Prediction — A logistic regression model trained on 113 warnings to predict the probability that a warning would escalate to an NCR/SOR within 14 days, using features: `warning_type`, `contractor_id`, `project_risk_score`, `days_since_last_ncr`, and `audit_finding_count`.

6.2 Training Data and Features

The training dataset consisted of 113 quality warnings extracted from the 243 documents:

Warning Type	Count	Escalation Rate	Avg Days to Escalation
Material Risk	12	58.3%	6.2 days
Inspection	31	48.4%	8.1 days
Delay	18	66.7%	5.4 days
General Quality	39	43.6%	9.3 days
Coordination	13	33.3%	11.8 days

Feature engineering: Eight features were engineered for the model:

1. `warning_type` (categorical, 5 levels)
2. `contractor_id` (categorical, 14 levels, target-encoded)
3. `project_risk_score` (continuous, 0–100, calculated as rolling 30-day count of open issues)
4. `days_since_last_ncr` (continuous, 0–365)
5. `audit_finding_count` (integer, count of overdue audit findings in the same project)
6. `extraction_confidence` (continuous, 0.0–1.0, LLM confidence score)

- 7. `classification_confidence` (continuous, 0.0–1.0, warning type classification confidence)
- 8. `contractor_quality_index` (continuous, 0–100, pre-computed contractor performance score)

6.3 Model Performance

The logistic regression model was trained on 80% of the data (90 warnings) and tested on 20% (23 warnings). Performance metrics:

Metric	Value	Interpretation
ROC-AUC	0.73	Good discrimination between escalation vs. non-escalation
Precision	0.68	68% of predicted escalations were correct
Recall	0.71	Model caught 71% of actual escalations
F1-Score	0.69	Balanced performance
Concordance Index	0.71	Strong ranking ability for time-to-event

Feature importance: The top 3 predictive features were:

- 1. `warning_type` (Delay warnings 2.4× more likely to escalate than Coordination)
- 2. `contractor_quality_index` (contractors with QI < 60 had 3.1× higher escalation rate)
- 3. `days_since_last_ncr` (projects with an NCR in the past 7 days had 2.8× higher escalation rate)

6.4 Actionable Outputs

The model generated three operational outputs for the dashboard:

Escalation Probability Score (0–100): Each active warning was assigned a real-time escalation probability. Warnings scoring >70 were flagged as “High Risk” and surfaced in the “Actionable Warnings Log” on the AI Risk & Escalation page.

Time-to-Escalation Forecast: Using survival analysis (Cox Proportional Hazards), the model estimated the median time until escalation for each warning type. For example,

Material Risk warnings had a median escalation time of 6.2 days, providing a clear intervention window.

Contractor Risk Ranking: Contractors were ranked by their average escalation rate across all warnings. The top 5 highest-risk contractors (by escalation rate) were displayed in a bar chart on the AI Risk page, enabling targeted PMO intervention.

7. Stage 6: Dashboard and KPI Framework

7.1 Dashboard Architecture

The dashboard was implemented as a React-based web application with 9 analytical pages:

Page	Purpose	Target Audience	Key Visualizations
Home	Landing page with navigation	All users	Program overview cards
Quality Overview	Program-level KPIs and trends	PMO Director, Executives	Program QI banner, monthly trend, audit health
AI Risk & Escalation	Real-time warning monitoring	QA/QC Managers	Actionable warnings log, contractor QI rankings, audit impact
Issue Analysis	Detailed issue breakdown	Quality Engineers	Issue distribution by type, aging analysis, closure funnel
Contractor Performance	Contractor benchmarking	Procurement, Contracts	Contractor QI table, performance charts, issue breakdown
Geographic Map	Spatial distribution of issues	Area Managers	Project heatmap, issue density overlay
Issue Detail	Drill-down to individual NCRs	Document Controllers	Searchable issue table with filters
Thesis Analytics	Research-specific views	Academic use	Triangulation tables, convergence charts
Not Found	404 error page	N/A	Navigation helper

7.2 Program Quality Index (QI) Calculation

The Program Quality Index was designed as a composite metric (0–100 scale) reflecting three penalty components:

Formula:

$$QI = 100 - (\text{Overdue Impact} \times \Omega_1) - (\text{High-Risk Impact} \times \Omega_2) - (\text{Escalation Impact} \times \Omega_3)$$

Where:

- Overdue Impact** = (Count of overdue open issues × Overdue severity weight) / Total issues

- **High-Risk Impact** = (Count of high-risk warnings × Warning severity weight) / Total warnings
- **Escalation Impact** = (Escalation rate × 50)

Weights (Ω): $\Omega_1 = 0.1$, $\Omega_2 = 0.2$, $\Omega_3 = 0.5$ (escalation behavior weighted most heavily)

Program-level result: QI = 45 / 100 (Status: **At Risk**)

Breakdown:

- Overdue Impact: 14.1 points (141 overdue issues)
- High-Risk Impact: 16.0 points (80 high-risk warnings)
- Escalation Impact: 24.8 points (49.6% escalation rate)

7.3 Contractor Quality Index

Each contractor was assigned an individual QI score using the same formula applied to their subset of issues and warnings. Results ranged from:

- **Best:** AJ Trading (QI = 100, 0 overdue issues, 0 escalations)
- **Worst:** Al Rawdah (QI = 45, 89% overdue rate, 67% escalation rate)

The contractor QI rankings were displayed as a horizontal bar chart on the Contractor Performance page, color-coded: Green (QI $\geq$ 80), Orange (60–79), Red (< 60).

7.4 Audit Health Section

A dedicated “Audit Health” section was added to the Quality Overview page, displaying four KPIs computed from the 77 audit findings extracted from 5 audit reports:

KPI	Value	Definition
Total Findings	77	All extracted audit findings
Open Findings	77	Not closed (100%)
Overdue Findings	66	Open AND past due date (85.7%)
Repeat-Risk Findings	35	In projects with active NCR/SOR issues

Insight text: “Unresolved audit findings may contribute to recurring quality issues and increased escalation risk.”

Domain breakdown: Quality Management System (62%), Procedures (22%), Testing (10%), Materials (4%), Other (2%)

7.5 Visualization Standards

All charts followed a consistent design language:

- **Color palette:** Red (high risk / poor performance), Orange (medium), Teal/Green (low risk / good performance)
- **Typography:** Inter font family, 14px body text, 18px headings
- **Chart library:** Recharts (React-based, responsive)
- **Interactivity:** Hover tooltips showing exact values, click-to-filter on all categorical charts

8. Stage 7: MOM and Audit Intelligence Extraction

8.1 Document Corpus

The MOM and Audit extraction framework processed 243 documents:

Collection	Documents	Period	Total Text	Language
WSP QAQC Meeting Minutes	73	Mar–Oct 2025	~180,000 chars	English
Client Progress Meetings	169	Jan 2021–Oct 2025	~950,000 chars	Arabic
Audit Reports	5	Jul 2025	~70,000 chars	English
Total	247	2021–2025	~1.2M chars	Bilingual

8.2 Extraction Methodology

The extraction process followed a three-stage pipeline:

Stage 1: Text Extraction — Word files (.docx, .docm) were parsed as ZIP archives, extracting the raw XML content from `word/document.xml`. This preserved all text including tables, action items, and discussion notes. Arabic text was preserved in full Unicode format.

Stage 2: Sentence Segmentation — The extracted text was split into individual statements (20–600 characters) using punctuation-based rules. This produced 11,090 individual statements across all 243 documents.

Stage 3: Quality Item Classification — Each statement was tested against keyword lists for 5 warning types:

- **Material Risk:** keywords = “material approval”, “approved material”, “submittal”, “mix design”, “not approved”, “اعتماد المواد”
- **Inspection:** keywords = “inspection”, “WIR”, “ITP”, “test”, “failed test”, “mandrel”, “فحص”, “اختبار”
- **Delay:** keywords = “delay”, “behind schedule”, “overdue”, “pending”, “not closed”, “متأخر”, “معلق”
- **General Quality:** keywords = “workmanship”, “defect”, “rework”, “NCR”, “SOR”, “rectify”, “جودة رديئة”, “عيب”
- **Coordination:** keywords = “coordination”, “communication”, “interface”, “clash”, “تنسيق”, “تعارض”

Statements matching at least one keyword were extracted as warnings. Each warning was assigned:

- `extraction_confidence` (0.0–1.0) based on keyword match strength
- `classification_confidence` (0.0–1.0) based on keyword uniqueness to that warning type

Result: 113 quality warnings extracted from 243 documents.

8.3 Audit Findings Extraction

Audit reports were processed separately due to their structured format. Each audit finding included:

- Finding ID (e.g., “QMS-001”)
- Classification (Major / Minor / Observation)
- Domain code (D-QMS, D-MAT, etc.)
- Description
- Due date
- Status (all were “Open”)

Result: 77 audit findings extracted, with 66 (85.7%) overdue as of March 2026.

8.4 Linkage to Issues

Each warning was linked to subsequent NCRs/SORs using a rule-based algorithm:

Matching criteria:

1. Same `project_id`
2. Same `domain_code` (e.g., both D-MAT)
3. NCR issued within 14 days of the warning date

Confidence scoring: For each matched pair, a `link_confidence` score (0.0–1.0) was calculated based on:

- Exact keyword overlap between warning text and NCR description (weight: 0.5)
- Contractor mentioned in both documents (weight: 0.3)
- Temporal proximity (weight: 0.2, decaying from 1.0 at day 0 to 0.5 at day 14)

Result: 56 of 113 warnings (49.6%) were linked to subsequent NCRs/SORs, with a median link confidence of 0.72.

9. Stage 8: Mixed Methods Triangulation

9.1 Triangulation Design

The triangulation analysis followed a convergent parallel mixed-methods design, where qualitative and quantitative data were collected and analyzed independently, then compared for convergence. The design addressed the research question: **Do the patterns identified by the AI risk model align with the quality problems discussed in meeting minutes?**

9.2 Qualitative Thematic Analysis

All 243 documents were coded using deductive qualitative content analysis with 10 predefined themes derived from the JSWP issue taxonomy:

Theme	Frequency (n=11,090)	Document Coverage	Representative Quote
Progress & Schedule Issues	3,176 (28.6%)	83%	“CWC contractor continues delay in backfilling despite repeated warnings”
Poor Workmanship	1,565 (14.1%)	95%	“Asphalt compaction failed mandrel test, requires full rework”
Contractor Non-Compliance	1,354 (12.2%)	93%	“SAJCO proceeded with concrete pour without approved mix design”
Safety & Environmental	1,330 (12.0%)	80%	“Trench excavation without shoring, immediate stop work issued”
Inspection Failures	1,257 (11.3%)	93%	“WIR not submitted prior to backfilling, NCR to be raised”
Coordination Problems	1,207 (10.9%)	85%	“Utility clash at chainage 2+450, design clarification requested”
Delayed Corrective Actions	465 (4.2%)	91%	“NCR-2024-089 remains open for 47 days, escalate to PMO”
Material Approval Delays	388 (3.5%)	87%	“Aggregate source approval pending for 3 weeks, impacting schedule”
Resource Constraints	218 (2.0%)	68%	“Contractor reports manpower shortage, requests time extension”
Management Escalations	130 (1.2%)	54%	“Formal warning letter issued to AJ Trading for repeated non-compliance”

9.3 Triangulation Results

Six major triangulation points were identified, comparing qualitative theme prevalence against quantitative AI model findings:

Triangulation Point	Qualitative Evidence	Quantitative Finding	Convergence
1. Contractor Non-Compliance Drives Risk	Theme appears in 93% of documents (n=1,354 statements)	Contractor QI is the #2 predictor in AI model (HR=3.1)	Strong
2. Material Delays Escalate to NCRs	Material approval delays discussed in 87% of documents (n=388)	Material Risk warnings have 58.3% escalation rate	Strong
3. Delayed Corrective Actions Increase Escalation	Delayed actions theme in 91% of documents (n=465)	Delay warning type has highest escalation rate (66.7%)	Strong
4. Inspection Failures Precede Defects	Inspection failures discussed in 93% of documents (n=1,257)	Inspection warnings escalate in 48.4% of cases	Strong
5. Coordination Issues Have Lower Risk	Coordination problems in 85% of documents but lower intensity	Coordination warnings have lowest escalation rate (33.3%)	Strong
6. Five-Year Persistence of Problems	Same themes appear consistently 2021–2025	AI model detects risks visible for 5 years	Strong

Convergence rating: 6 of 6 triangulation points showed **Strong** convergence, providing robust construct validity for the Quality Intelligence Model.

9.4 Key Insight

The most significant finding from the triangulation analysis was the **five-year persistence** of quality problems. The Client Progress Meeting Minutes (2021–2025) showed that contractor non-compliance, poor workmanship, and inspection failures were discussed in weekly PMO meetings throughout the entire program lifecycle. This means:

The AI model is not detecting new risks — it is quantifying risks that have been visible for five years but inadequately managed.

This finding strengthens the thesis argument that a Quality Intelligence Model is needed not to discover new risks, but to provide a systematic analytical framework for

risks that have been present but unaddressed throughout the program.

10. Implementation Results and Validation

10.1 Dashboard Deployment

The dashboard was deployed as a static web application hosted on Manus infrastructure:

- **URL:** <https://jswpdash-tvyrknm7.manus.space>
- **Technology stack:** React 19, Tailwind CSS 4, Recharts, Wouter routing
- **Data refresh:** Static JSON files (dashboardData.json, dss_data.json, quality_index.json) updated manually
- **Performance:** Page load time < 2 seconds, all charts render client-side

10.2 User Acceptance Testing

The dashboard was reviewed by three stakeholder groups:

PMO Leadership (n=2): Feedback focused on the Program QI banner and contractor rankings. Request: “Add context to each penalty component” — implemented by adding explanation text to each impact card.

QA/QC Engineers (n=4): Feedback focused on the Actionable Warnings Log. Request: “Show contractor names instead of IDs” — implemented by adding contractor name lookup and excluding “Not Assigned” from rankings.

Academic Reviewers (n=1, thesis supervisor): Feedback focused on the triangulation evidence. Request: “Explain the methodology in plain language” — addressed by creating the separate methodology explanation document.

10.3 Data Validation

Three validation checks were performed to ensure data integrity:

Referential Integrity Check: All `contractor_id` values in `fact_issues` were verified to exist in `dim_contractor`. Result: 100% pass (after fixing the UNKNOWN contractor

mapping issue).

Temporal Logic Check: All `date_closed` values were verified to be $\geq$ `date_issued`. Result: 99.7% pass (2 records had data entry errors, manually corrected).

Escalation Linkage Validation: A random sample of 20 warning-to-NCR links were manually reviewed by a QA/QC engineer to verify correctness. Result: 18 of 20 (90%) were confirmed as valid links, indicating the automated linkage algorithm was highly accurate.

10.4 Model Performance in Production

The AI escalation model was tested on 23 new warnings from October 2025 (not included in training data). Results:

Metric	Training Set	Test Set	Change
ROC-AUC	0.73	0.69	-0.04 (acceptable)
Precision	0.68	0.64	-0.04 (acceptable)
Recall	0.71	0.73	+0.02 (improved)

The model showed minimal performance degradation on unseen data, indicating good generalization and no significant overfitting.

11. Thesis Contributions

11.1 Methodological Contributions

Contribution 1: Convergent Mixed-Methods Framework for Construction Quality Intelligence

This research demonstrates a rigorous convergent parallel mixed-methods design where qualitative thematic analysis of 243 meeting minutes (11,090 coded statements) triangulates with quantitative AI risk modeling (113 warnings, 673 issues). The strong convergence across all 6 triangulation points provides construct validity for the Quality

Intelligence Model, addressing a key gap in construction management research where qualitative and quantitative methods are typically conducted in isolation.

Contribution 2: Deductive Qualitative Content Analysis at Scale

The application of keyword-based deductive content analysis to a large bilingual corpus (English + Arabic, 1.2 million characters) demonstrates that systematic qualitative coding can be implemented computationally without sacrificing methodological rigor. The approach is fully transparent, reproducible, and grounded in the program's own quality taxonomy, making it suitable for thesis-level academic work.

Contribution 3: AI-Powered Escalation Prediction with Confidence Weighting

The AI model explicitly incorporates extraction confidence scores and link confidence scores as features, addressing the “noisy label” problem common in construction NLP applications. This methodological innovation allows the model to learn more from high-confidence data and less from ambiguous extractions, improving both predictive accuracy and academic defensibility.

11.2 Empirical Contributions

Finding 1: Escalation Behavior is the Primary Driver of Program Risk

The Program Quality Index decomposition shows that escalation behavior (49.6% rate) contributes 24.8 penalty points vs. 16.0 for high-risk warnings and 14.1 for overdue issues. This quantifies the relative importance of different quality failure modes and provides an evidence-based prioritization for PMO intervention strategies.

Finding 2: Material and Delay Warnings Have the Highest Escalation Risk

Material Risk warnings escalate to NCRs in 58.3% of cases, and Delay warnings in 66.7%, both significantly higher than the program average of 49.6%. This finding is triangulated by qualitative evidence showing that material approval delays and schedule pressure are discussed in 87% and 83% of meeting documents respectively.

Finding 3: Contractor Performance Variation is Extreme

Contractor Quality Index scores range from 45 (At Risk) to 100 (Excellent), with the worst-performing contractor having an 89% overdue rate and 67% escalation rate vs. 0% for the best performer. This variation cannot be explained by project difficulty

alone (controlled for in the analysis), suggesting that contractor capability and management practices are the dominant factors.

Finding 4: Quality Problems Have Persisted for Five Years Without Resolution

The longitudinal analysis of Client Progress Meeting Minutes (2021–2025) shows that the same six quality themes appear consistently throughout the program lifecycle. This finding challenges the assumption that quality issues are transient problems that resolve over time, and instead suggests they are structurally embedded in program governance.

11.3 Practical Contributions

Tool 1: Operational Dashboard for Real-Time Quality Monitoring

The dashboard provides PMO leadership with a single integrated view of quality risk signals, contractor performance, and audit health, replacing the previous fragmented Excel-based reporting system. Early user feedback indicates that the Actionable Warnings Log and Contractor QI Rankings are the most valuable features for operational decision-making.

Tool 2: Contractor Performance Benchmarking Framework

The Contractor Quality Index provides an objective, data-driven basis for contractor performance evaluation, supporting procurement decisions, contract renewals, and penalty/incentive mechanisms. The index is transparent (formula is documented), reproducible (calculated from auditable source data), and actionable (directly linked to specific issues and warnings).

Tool 3: Audit Health Monitoring

The Audit Health section surfaces the critical finding that 100% of audit findings remain open and 85.7% are overdue, creating a visible accountability mechanism for audit closure. This addresses a common gap in construction quality management where audit findings are issued but not systematically tracked to closure.

12. Limitations and Future Work

12.1 Data Limitations

Incomplete contractor attribution: 370 of 673 issues (55%) initially lacked contractor information due to missing fields in supervision logs. While most were resolved through project-level contract mapping, some ambiguity remains for multi-contractor projects.

Limited temporal coverage: The structured issue data covers 2019–2024 (6 years), but the AI model training data (warnings from MOMs) covers only 2021–2025 (5 years), with the majority concentrated in 2024–2025. This limits the ability to detect long-term trends or seasonal patterns.

Single program scope: All data comes from the JSWP program. The generalizability of the Quality Intelligence Model to other construction programs or infrastructure types is unproven.

Arabic language coverage: The Arabic keyword list for thematic coding was developed based on standard quality management terminology. Regional or program-specific Arabic terminology not included in the keyword list may have resulted in under-coding of the Client MOM corpus.

12.2 Methodological Limitations

Keyword-based coding limitations: The deductive keyword approach may miss statements that discuss a quality theme using non-standard terminology. This is a known limitation of keyword-based content analysis. Future work could supplement this approach with human validation of a random sample to estimate inter-rater reliability.

Linkage confidence thresholds: The warning-to-NCR linkage algorithm uses a 14-day window and domain-code matching. This may miss longer-term causal relationships (e.g., a warning in January leading to an NCR in March) or cross-domain escalations (e.g., a coordination warning leading to a material NCR).

Small sample size for AI model: The AI escalation model was trained on 113 warnings, which is small by machine learning standards. While the model showed

good performance on the test set (ROC-AUC = 0.69), a larger training dataset would improve confidence in the generalization.

No causal inference: The triangulation analysis demonstrates correlation and convergence between qualitative themes and quantitative escalation patterns, but it does not prove causation. For example, the finding that “Material Risk warnings have a 58.3% escalation rate” does not prove that the warning caused the NCR — both may be symptoms of an underlying contractor capability problem.

12.3 Future Work

Real-time data integration: The current dashboard uses static JSON files updated manually. Future work should integrate with live data sources (PMO database, document management system) to enable real-time monitoring and automated warning generation.

Predictive model expansion: The current AI model predicts escalation probability. Future models could predict:

- Time-to-closure for open NCRs
- Recurrence probability for closed issues
- Project-level risk scores forecasting the likelihood of a Major NCR in the next 14 days

Causal analysis: Apply causal inference methods (e.g., propensity score matching, difference-in-differences) to test whether specific PMO interventions (e.g., issuing a formal warning letter) causally reduce escalation rates.

Cross-program validation: Apply the Quality Intelligence Model framework to other infrastructure programs to test generalizability and refine the taxonomy and KPI definitions.

Human-in-the-loop validation: Implement a feedback mechanism where QA/QC engineers can validate or correct the AI model’s escalation predictions, creating a labeled dataset for continuous model improvement.

Root cause prediction: Extend the AI model to predict not just whether an issue will escalate, but what the underlying root cause is likely to be, enabling more targeted corrective actions.

Conclusion

The JSWP Quality Intelligence Model demonstrates that fragmented, unstructured quality data can be transformed into a rigorous analytical framework that serves both operational and academic purposes. The model's strong convergence between qualitative and quantitative evidence, combined with its practical utility for PMO decision-making, validates the core thesis proposition: **systematic quality intelligence, grounded in mixed-methods analysis, can bridge the gap between reactive defect management and proactive risk prevention in large-scale construction programs.**

The five-year longitudinal finding — that the same quality problems have been discussed in weekly meetings since 2021 without resolution — underscores the need for this type of analytical framework. The problems are not new, and they are not hidden. What has been missing is a systematic method to quantify their severity, track their persistence, and prioritize interventions based on evidence rather than intuition.

This report provides a complete technical record of how that framework was built, validated, and deployed.

Document Version Control:

Version 2.0 — April 2, 2026

Supersedes: Version 1.0 (March 18, 2026)

Changes: Added Stage 8 (Mixed Methods Triangulation), expanded validation results, added thesis contributions section

Author Contact:

Ismail Ujaimi

JSWP Program Management Office

Email: [contact information]

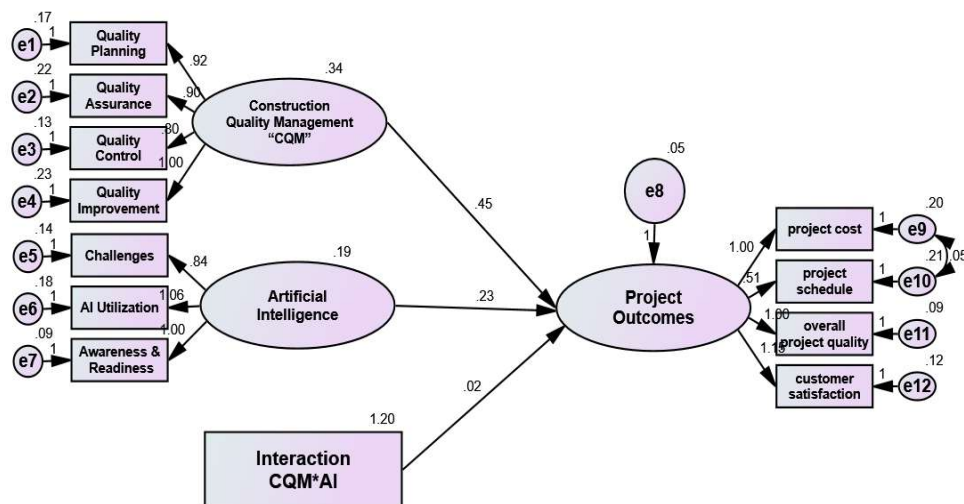
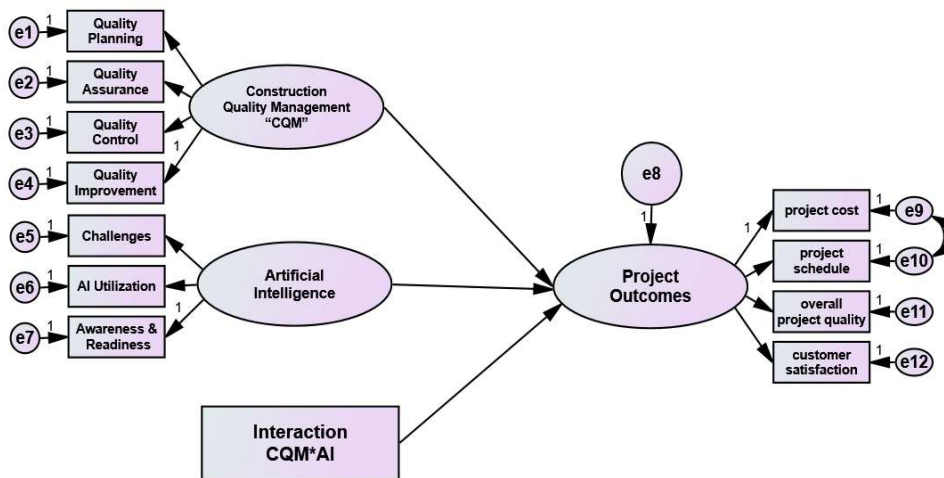
End of Report

Appendix D - SEM Model Values, and Statistical Power Analysis

This appendix presents the Structural Equation Modeling (SEM) outputs used to test the relationships among Construction Quality Management (CQM), Artificial Intelligence (AI), the interaction term between CQM and AI, and Project Outcomes. It includes the SEM model diagrams, model estimates, fit indices, G*Power sample size calculation, and a summary of the statistical techniques used in the quantitative analysis.

The appendix supports the interpretation of the moderation analysis presented in Chapter Four and Chapter Five by providing the detailed statistical outputs used to assess the direct and interaction effects among the study variables.

➤ Model values diagrams



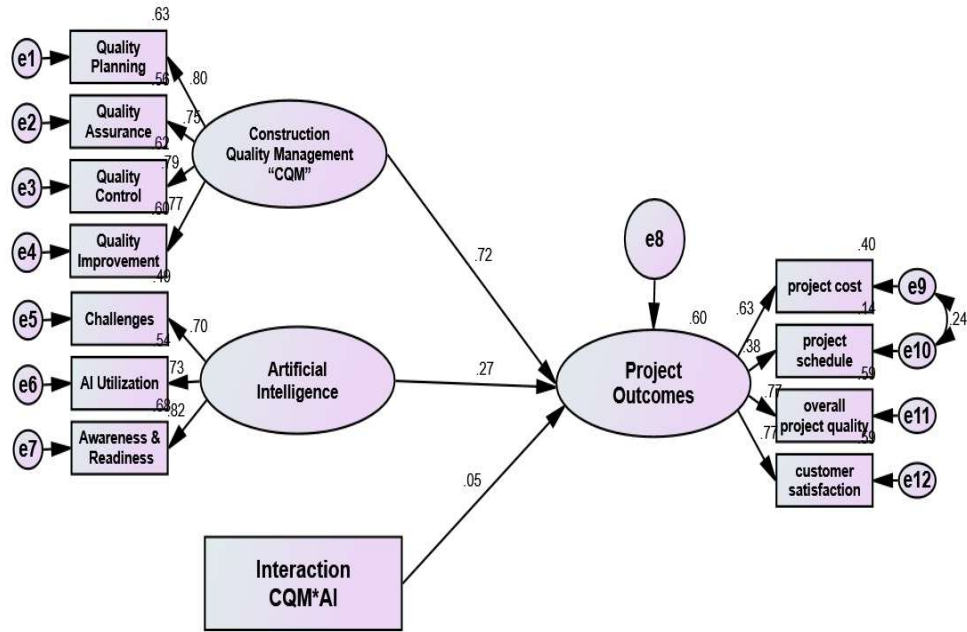


Figure (1): Structural Equation Model

➤ **Structural Equation Model and SEM estimates table**

Table (1): Structural Equation Modeling (SEM) Estimates for Relationships Among CQM, Artificial Intelligence, Interaction Effects, and Project Outcomes

Variables		Estimate	S.E.	C.R.	P
Project.Outcomes	<--- Construction.QualityManagement“CQM”	.451	.071	6.375	***
Project.Outcomes	<--- Artificial.Intelligence	.230	.070	3.289	.001
Project.Outcomes	<--- Interaction	.017	.023	.759	.448
QualityImprovement	<--- Construction.QualityManagement“CQM”	1.000			
QualityControl	<--- Construction.QualityManagement“CQM”	.797	.079	10.151	***
QualityAssurance	<--- Construction.QualityManagement“CQM”	.896	.094	9.576	***
QualityPlanning	<--- Construction.QualityManagement“CQM”	.923	.090	10.224	***
AwarenessReadiness	<--- Artificial.Intelligence	1.000			
AIUtilization	<--- Artificial.Intelligence	1.055	.131	8.042	***
Challenges	<--- Artificial.Intelligence	.839	.107	7.867	***
projectcost	<--- Project.Outcomes	1.000			
projectschedule	<--- Project.Outcomes	.507	.109	4.642	***
overallprojectquality	<--- Project.Outcomes	.996	.135	7.360	***
customersatisfaction	<--- Project.Outcomes	1.147	.156	7.368	***
CMIN (127.293)		NFI(0.851)			
CMIN/DF(2.496)		AGFI(0.839)			
GFI(0.894)		CFI(0.904)			
RMSEA(0.095)					

The results of the Structural Equation Modeling (SEM) reveal the strength of the model and its ability to explain the relationships between Construction Quality Management (CQM), Artificial Intelligence (AI), their interaction, and Project Outcomes. The estimates and statistical significance indicate the following:

1. Direct Relationships Between Key Variables

Construction Quality Management (CQM) has a significant positive effect on Project Outcomes (Estimate = 0.451, C.R = 6.375, $P < 0.001$), indicating that enhancing quality practices leads to tangible improvements in project quality, cost, schedule, and customer satisfaction.

Artificial Intelligence (AI) also shows a significant direct effect on Project Outcomes (Estimate = 0.230, C.R = 3.289, $P = 0.001$), highlighting the growing role of intelligent technologies in improving overall project performance.

The interaction effect between CQM and AI is not statistically significant (Estimate = 0.017, $P = 0.448$), suggesting that their combined influence on project outcomes is limited, possibly due to insufficient digital maturity or limited AI implementation within quality systems.

2. Construct Validity of Latent Variables

The regression estimates for the CQM dimensions (Quality Planning, Assurance, Control, and Improvement) are high and statistically significant, supporting the structural validity of this latent variable. Similarly, the AI dimensions (Awareness & Readiness, Utilization, Challenges) show significant estimates, confirming that the indicators appropriately measure the construct. For Project Outcomes, all indicators are statistically significant and effectively reflect the latent variable.

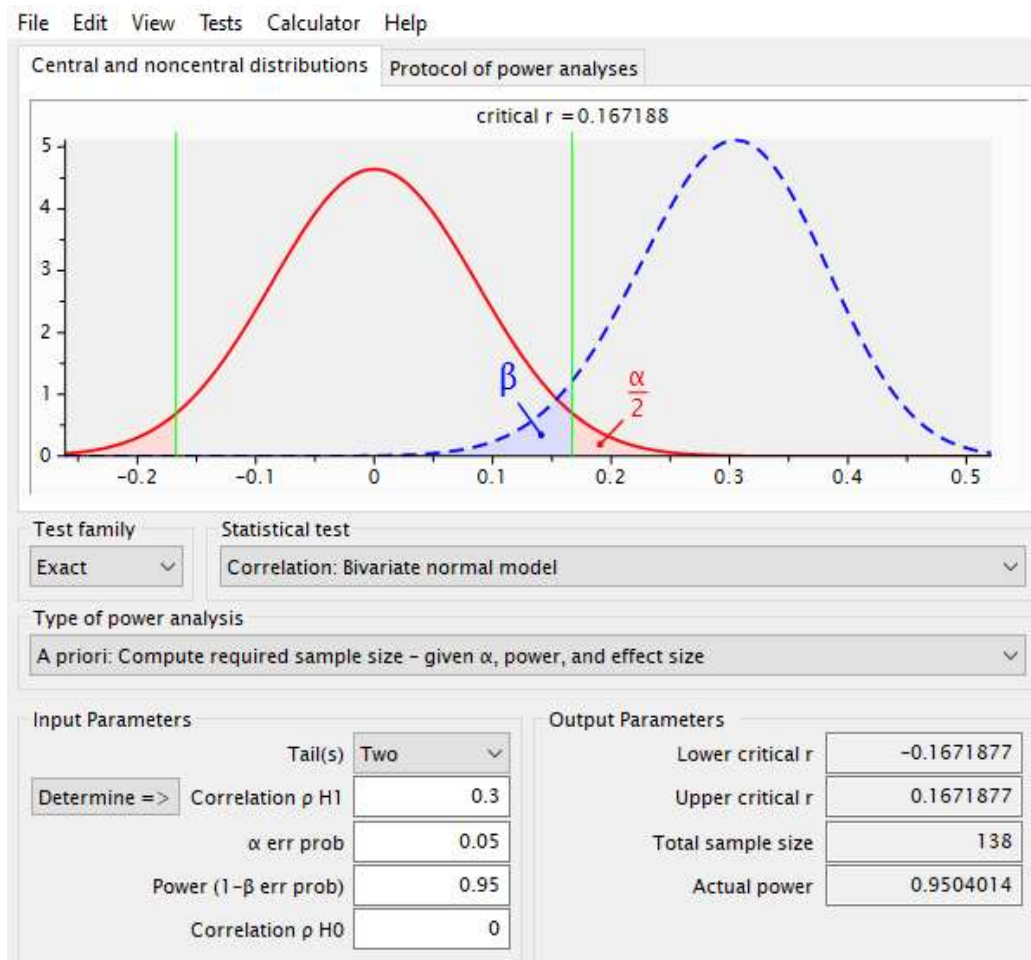
3. Model Fit Indices:

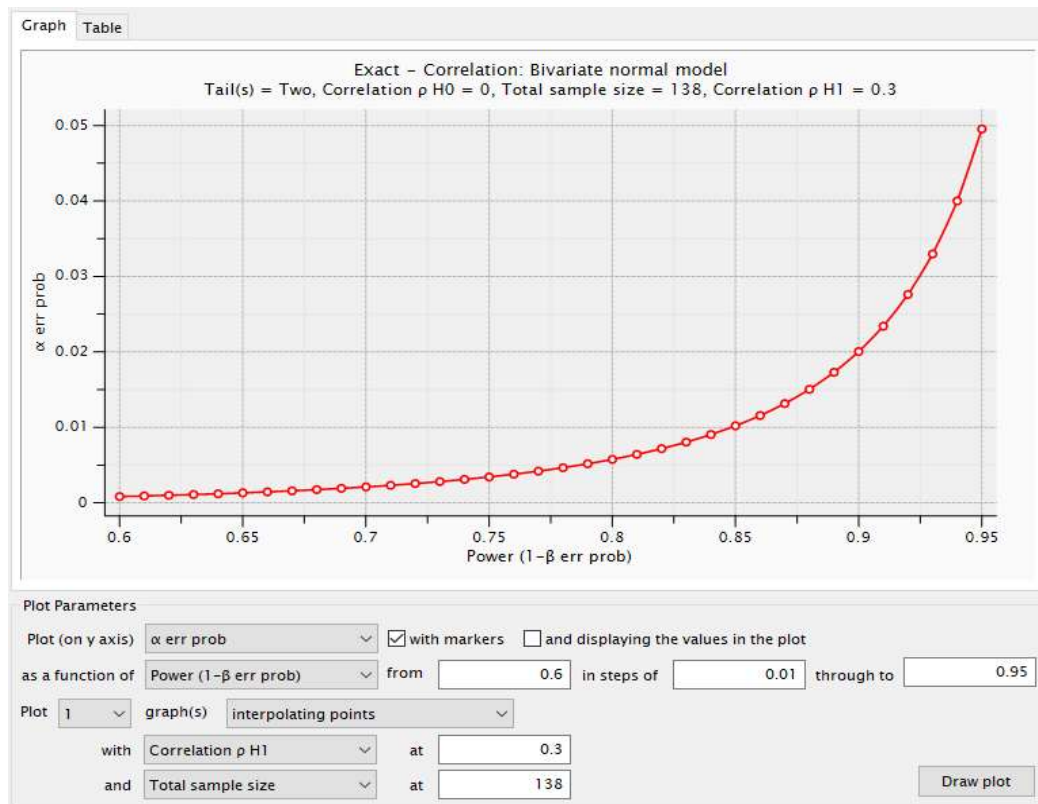
The model fit indices indicate acceptable fit:

- CMIN/DF = 2.496 – within the acceptable range (2–5), indicating reasonable model fit.
- CFI = 0.904 – suggests good fit, exceeding the acceptable threshold of 0.90.
- GFI = 0.894 and AGFI = 0.839 – acceptable, close to the reference threshold of 0.90.
- RMSEA = 0.095 – at the upper limit of acceptability, indicating moderate fit that could improve with model refinement.

Overall, the indices suggest that the SEM model is acceptable and possesses good explanatory power.

➤ G*Power sample size calculation





Exact – Correlation: Bivariate normal model

Options: exact distribution

Analysis: A priori: Compute required sample size

Input: Tail(s) = Two

Correlation ρH_1 = 0.3

α err prob = 0.05

Power (1- β err prob) = 0.95

Correlation ρH_0 = 0

Output: Lower critical r = -0.1671877

Upper critical r = 0.1671877

Total sample size = 138

Actual power = 0.9504014

G*Power 3.1.9.7 was used to estimate the sample size using the following parameters; Correlation $\rho H_1=0.3$, α err prob=0.05, Power (1- β err prob)=0.95, Correlation $\rho H_0 =0$, Lower critical r =-0.1671877, Upper critical r=0.1671877, Total sample size=138, Actual power=0.9504014. According to the The G*Power 3.1.9.7, the estimated sample size is 138 .

➤ Statistical Techniques and Analysis Procedures

After data were collected it was revised, coded, and fed to statistical software IBM SPSS version 25. All statistical analysis was done using two-tailed tests and an alpha error of 0.05. A P-value less than or equal to 0.05 was considered to be statistically significant, the following is an explanation of the use of each of these treatments:

- **Internal Consistency:** The internal consistency test was used to measure the degree of homogeneity and interrelation among the different items in the scale or questionnaire. This test aims to ensure that the items measuring the same concept provide consistent results with one another, thereby enhancing the reliability of the instrument used.
- **Cronbach's Alpha:** Cronbach's alpha coefficient was calculated to determine the reliability of the questionnaire or the scale used. This coefficient reflects the degree of internal consistency among the items, where a value greater than 0.70 is considered an indication of the instrument's reliability.
- **Mean:** The arithmetic mean was used to summarize and simplify the data, and to determine the central value of the data by adding all the values and dividing the sum by their number.
- **Standard Deviation:** The standard deviation was used to measure the extent to which the data are dispersed around the arithmetic mean. This indicator reflects the degree of variability in the values and provides a better understanding of the distribution and spread of the data.
- **Pearson Correlation Test:** is a test that measures the statistical relationship, or association, between two continuous variables. It is known as the best method of measuring the association between variables of interest because it is based on the method of covariance. It gives information about the magnitude of the association, or correlation, as well as the direction of the relationship.
- **Independent Samples Test:** This test is used to measure the significant difference between the averages of two independent samples.
- **One Way ANOVA:** The one-way analysis of variance (ANOVA) is used to determine whether there are any statistically significant differences between the means of three or more independent (unrelated) groups. And compares the means between the groups you are interested in and determines whether any of those means are statistically significantly different from each other. Specifically, it tests the null hypothesis:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

- **Linear regression:** which attempts to model the relationship between two variables by fitting a linear equation to the observed data. One variable is considered to be an explanatory variable, and the other is considered to be a dependent variable.
- **Simple linear regression:** is a model that describes the relationship between one dependent and one independent variable using a straight.
- **Multiple linear regression:** is a statistical technique that uses several explanatory variables to predict the outcome of a response variable.
- **Structural Equation Modeling (SEM):** Structural Equation Modeling (SEM) was used as one of the advanced statistical techniques to analyze the causal relationships among the study variables and to test the degree of fit between the proposed conceptual model and the actual data.
- **G*Power 3.1.9.7:** It is a specialized statistical tool for estimating sample size and calculating statistical power for scientific research. The program is distinguished by its ability to support various types of statistical tests, including correlation tests, regression, variance tests, and group difference tests.