

BUBBLES

AI + Creativity + ROI

SOCIAL BUBBLES IT SERVICES L.L.C · DUBAI

CONSTRUCTION TAKEOFF SUITE

Detailed System & Architecture Document

AI-assisted tender analysis, quantity takeoff, materials and procurement — every quantity measured, every risk surfaced, every number traceable to its source drawing.

DOCUMENT	First-touch technical & commercial compendium
SCOPE	Eight lifecycle stages · engine architecture · security · expected results · pilot · FAQ
REVISION	REV A · AUG 2026
STATUS	CONFIDENTIAL — for the addressed recipient

ABOUT THIS DOCUMENT

How to read this document

This compendium expands the first-touch presentation of the Construction Takeoff Suite into a reference you can circulate internally: to the estimating desk, to management, and to IT. It is organised so that each audience can enter at a different door and still meet in the middle.

If you are...	Start with...	Then read...
The MD / owner	Part I (the problem) and Part VI (impact & expected results)	Part VII (the pilot) — the zero-risk way to test every claim
The chief estimator / QS	Part III (the eight stages, two layers each)	Part IV (architecture & the five trust layers) — why the numbers hold
IT / compliance	Part V (security, confidentiality, deployment)	Part IV §4.6 (the LLM trust boundary) and the FAQ security section
In a hurry	The executive summary overleaf	Part VIII (FAQ) — most objections are answered there

Throughout, **illustrations carry the argument** and the text carries the detail. Every claim that can be tied to a measured, delivered engagement is tied to one; forward-looking targets are labelled as such. Real project examples are drawn from a live pilot on a G+2P+8 residential tower tender (1,557 files) processed end-to-end.

ONE SENTENCE, IF YOU KEEP ONLY ONE

The suite follows the life of a tender through eight stages on one shared evidence spine — **AI reads and labels, deterministic code counts and computes, a qualified QS signs** — so estimating capacity stops deciding which tenders you can attempt, and every number in the bill can be defended line by line.

CONTENTS

Table of contents

How to read this document 2

Table of contents 3

Executive summary 6

1. The tender, as it is run today 7

 1.1 Three weeks, one deadline, zero slack 7

 1.2 What a bid actually costs 7

 1.3 The two silent failure modes 7

2. One spine, eight stages 9

 2.1 The evidence spine 9

 2.2 The division of labour 9

 2.3 A system that learns your business 9

01. Intake & Tender Study 11

 What happens, in detail 11

Stage 01 — deliverables & operational benefit 13

 Operational benefit 13

02. Red Flags & Clarifications 14

 What happens, in detail 14

Stage 02 — deliverables & operational benefit 15

 Operational benefit 15

03. Bid / No-Bid Decision Room 16

 What happens, in detail 16

Stage 03 — deliverables & operational benefit 17

 Operational benefit 17

04. The Bill of Quantities 18

 What happens, in detail 18

Stage 04 — deliverables & operational benefit 19

 Operational benefit 19

05. The Bill of Materials 20

 What happens, in detail 20

Stage 05 — deliverables & operational benefit 21

 Operational benefit 21

06. Addendum & Revision Watch 22

What happens, in detail 22

Stage 06 — deliverables & operational benefit 23

Operational benefit 23

07. Procurement & Subcontracting 24

What happens, in detail 24

Stage 07 — deliverables & operational benefit 25

Operational benefit 25

08. Award & Beyond 26

What happens, in detail 26

Deliverables & operational benefit 26

4. How the engine actually works 27

4.1 The whole route on one board 27

4.2 The pipeline, stage by stage 28

4.3 The project graph and the provenance contract 29

4.4 Where documents disagree: the precedence ladder 29

4.5 Exactly where AI is allowed — and where it is not 30

4.6 Five structural layers against hallucination 30

4.7 Auditing the engine — and auditing the auditor 31

4.8 Accuracy as an institution: golden corpora 31

4.9 How measurement standards become code 32

What a rule pack contains 32

4.10 The seven human gates 32

5. Confidentiality by architecture 34

5.1 Local-first processing 34

5.2 Integrity and isolation 34

5.3 Data residency, your way 34

5.4 What the AI models actually see 34

6. Where the weeks go — and come back 35

6.1 Stage-by-stage time impact 35

6.2 The cumulative picture 35

6.3 Team activity tracking 35

7. Expected results — and the convergence effect 36

7.1 Why the numbers improve with every tender 36

8. Five ways this outbids the field 37

9. Prove it on your own project — NDA first 38

9.1 The pilot, step by step

38

9.2 What the pilot is for

38

Annex A — what the deliverables actually look like

39

A.1 A finding, dissected

39

A.2 A bill line, dissected

40

A.3 The audit certificate

40

Annex B — from NDA to production, week by week

41

10. Frequently asked questions

42

Accuracy & trust

42

Data & security

42

People & workflow

43

Commercial & engagement

43

Technical

44

Adoption & operations

44

Annex C — glossary of terms

46

11. The proposition, restated

48

EXECUTIVE SUMMARY

Executive summary

Construction tendering is a race run on manual measurement. Between the day a tender lands and the day the bid is submitted, an estimating team must read thousands of pages, reconcile drawings against schedules and specifications, quantify every trade, chase vendor quotes and survive late addenda — all inside a 3–6 week window. The work is skilled, expensive, and largely invisible until something goes wrong: a quantity mis-measured with confidence, or a scope item missed entirely. Both failure modes are silent at submission and expensive at execution.

The **Construction Takeoff Suite** by Social Bubbles IT Services LLC is a drawing-to-bill platform that automates the measurable work of tendering while keeping professional judgement — and sign-off — with your qualified quantity surveyors. It follows the tender through eight stages: intake & study, red flags & clarifications, bid/no-bid decision support, the Bill of Quantities, the Bill of Materials, addendum & revision watch, procurement & subcontracting, and award & beyond. Every stage writes to a single **evidence spine**: every fact carries the file, sheet, revision and method it came from.

Three design principles separate the suite from generic “AI for construction” tools:

- **AI never counts, measures or prices.** Models read and label documents; deterministic code performs every measurement and calculation through re-executable formulas. Five structural layers (Part IV) close every route by which an AI hallucination could reach a number.
- **Honesty is enforced, not promised.** Anything the engine cannot measure is flagged with the reason — the bill is permitted to be incomplete, never wrong. An independent audit pass re-derives the numbers before anything is issued.
- **The system learns your business.** Consultant profiles, measurement rule packs, consumption norms and a growing rate library make every processed tender cheaper, faster and more accurate than the last — post-calibration accuracy converges to $\pm 2\text{--}3\%$, with manual review confined to a flagged $\sim 3\%$ of lines and shrinking further as the corpus grows.

1,557	15/15	5,934	212	165	0
FILES INGESTED	BILL SECTIONS	ITEMS MEASURED	ISSUES SURFACED	QUERIES DRAFTED	FABRICATED NUMBERS

Field proof: a complete live tender processed end-to-end under unfamiliar consultant conventions, with benchmark trades within $\pm 3\%$ (concrete & formwork) and $\pm 5\%$ (rebar) of a client’s own issued bill.

The commercial proposition is equally simple: measured on the pilot, the tender study compresses from 4–5 days to one, full takeoff from weeks to days, the late-addendum response from days to about an hour, and the RFQ cycle from weeks to days — a cumulative cycle-time reduction of roughly 70–75%. The freed hours return to judgement: pricing strategy, risk decisions, and negotiation. The engagement begins with a **no-obligation NDA-first pilot** on one of your own live projects (Part VII), so every claim in this document can be verified against your own numbers before any commercial commitment.

PART I · THE PROBLEM

1. The tender, as it is run today

Everything the suite does is aimed at a three-to-six-week window that has barely changed in fifty years, even as project documentation has grown tenfold.

1.1 Three weeks, one deadline, zero slack

A typical invited tender arrives as a multi-gigabyte dump: five or more volumes covering conditions of contract, specifications, an unpriced Bill of Quantities, hundreds of drawings across architecture, structure and MEP, and a soil investigation report — followed, in the weeks after, by addenda, clarification responses and revised drawings that quietly supersede what was first issued. From day zero, five clocks start ticking at once:

- **The study clock.** Someone must discover what has actually been received, what is missing, and what the commercial conditions really say — typically the first 4–5 working days of senior time.
- **The query clock.** Questions to the consultant must be lodged before a fixed deadline, usually in week two. Miss it, and every ambiguity becomes your pricing risk.
- **The measurement clock.** Quantity takeoff across all trades — the bulk of the estimating effort, typically two to three weeks of desk time.
- **The procurement clock.** Subcontractor and supplier quotes gate the pricing; enquiries must go out early enough for responses to come back and be compared.
- **The addendum clock.** Late revisions land in week three — precisely when there is no time left to re-measure what they change.

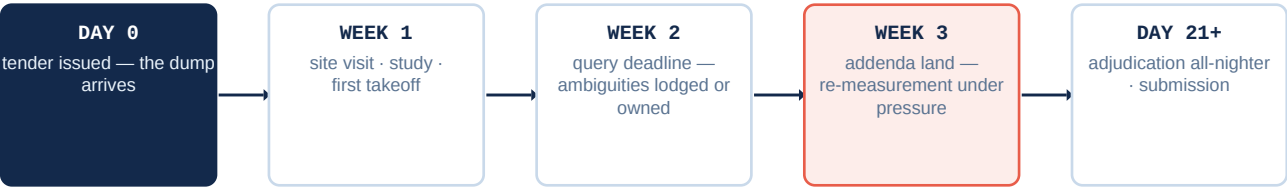


Figure 1.1 — the compressed tender window; the red box is where late addenda meet zero remaining slack.

1.2 What a bid actually costs

Industry rules of thumb put the cost of preparing a serious tender at **0.25–1% of contract value** — spent win or lose. For a mid-size contractor bidding ten tenders to win two or three, estimating is one of the largest unbilled cost centres in the business. Worse, the constraint is not money but **capacity**: the size of the estimating team decides how many tenders can even be attempted, which means the real cost of manual takeoff includes every winnable tender that was declined for bandwidth.

0.25–1%	3–6 wks	2–3 wks	0 hrs
OF CONTRACT VALUE PER BID	COMPRESSED WINDOW	OF THAT IS MEASURING	LEFT FOR JUDGEMENT

1.3 The two silent failure modes

When tenders lose money, the loss almost always traces back to one of two failure classes — both invisible on submission day, both discovered only when the price is already contractually binding.

Failure mode	What it is	Real examples (live tenders)
Silent quantity error	A number exists but is wrong: a misread unit, a wrong storey height, a missed deduction — priced with full confidence because nothing looked wrong.	A CAD schedule drawn in inches inside a metric set — a ×25.4 error for any tool or person who trusted it. A 0.4 m storey-height assumption moving blockwork ±12%.
Silent scope error	A number is missing entirely: a referenced document never issued, or two issued documents that disagree, so scope is priced off the wrong one.	Window family W-03: drawn 6–12 units on plans, 18–40 in the schedule. Pricing preambles PP/1–28 referenced in the bill index but never issued with the tender.

Manual review does not reliably catch either class, for a structural reason: the reviewer samples, and the errors hide in the unsampled majority. A senior QS checking a 6,000-item takeoff under deadline pressure re-measures perhaps two or three percent of it — chosen by instinct. Instinct is good; it is not exhaustive. The suite’s answer (Part III) is to check **everything**, deterministically, and to hand the human reviewer only the lines that deserve judgement.

THE COST, STATED PLAINLY
Both failure modes convert directly into margin: rework, disputes, unbudgeted materials, and claims fought without evidence. A single uncaught drawn-vs-scheduled conflict on one window family can exceed the cost of automating the entire tender.

PART II · THE SUITE AT A GLANCE

2. One spine, eight stages

The suite is not a single tool but a pipeline that follows the tender through its whole life — from the raw dump to final account — with one non-negotiable rule: every fact carries its evidence.

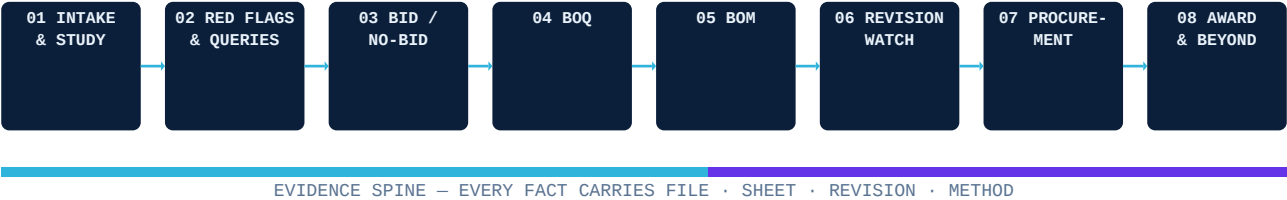


Figure 2.1 – the eight lifecycle stages over the shared evidence spine.

2.1 The evidence spine

Every sheet, schedule row, symbol, measurement, finding and bill line in the system carries a **provenance record**: which file it came from, which sheet and revision, where on the sheet, and by what method it was extracted. This is what makes the suite auditable rather than merely convincing: any number in any deliverable can be walked backwards to the exact drawing location that produced it. Provenance is also what makes incremental recomputation possible — when a revision lands, the engine knows precisely which downstream results cite the superseded sheet and must re-run (Stage 06).

2.2 The division of labour

Actor	Does	Never does
AI models	Read and label: classify sheets, map CAD layers to meanings, normalise messy schedule tables into typed rows, phrase descriptions, draft query letters from structured findings.	Count, measure, calculate, price, or invent a value. Every numeric output is guarded: a number that does not exist in the input is rejected automatically.
Deterministic code	Everything numeric: geometry measurement from CAD vectors, ~40 named validation checks, quantity formulas, BOM explosion, reconciliations, delta bills.	Interpret ambiguity. Where documents conflict, code applies a codified precedence ladder and flags the line — it does not silently choose.
Your QS	Judge: confirm layer mappings once per consultant, triage findings, approve assumptions, sign the bill. Review effort is concentrated on the flagged ~3% of lines.	Re-measure the other 97%. The engine's evidence makes spot-checking a choice, not a necessity.

2.3 A system that learns your business

The suite is **self-improving by design**, and the mechanism is concrete rather than rhetorical. Four assets accumulate with every processed tender: **consultant profiles** (title-block layouts, sheet-numbering schemes, layer conventions — learned once per consultant, reused forever), **measurement rule packs** (POMI/NRM2 plus your own amendments, versioned and citable), **consumption norms and the rate library** (your materials logic and every vendor quote ever received), and **the accuracy benchmark corpus** (every tender with a known-good human bill becomes a regression target the engine must keep matching). The practical consequence appears throughout this document: onboarding cost falls per

project, accuracy converges toward the $\pm 2\text{--}3\%$ band, and the manual-review share shrinks below its initial $\sim 3\%$ as the system adapts to your formats, your consultants and your trades.

PART III · STAGE 01

01. Intake & Tender Study

“Send the dump exactly as it arrived.” The first stage turns a chaotic multi-gigabyte delivery into a governed, complete, queryable set — the same day it lands.

What happens, in detail

The engine accepts the tender exactly as issued — nested ZIPs and RARs, mixed PDFs and CAD, Excel schedules, Word specifications, scanned addenda, junk files and all. No pre-sorting, no renaming, no “please structure your folders first”. Intake then proceeds through four deterministic passes:

- **Unpack & fingerprint.** Recursive extraction of every archive (including archives inside archives); every file SHA-256 hashed; exact duplicates detected and collapsed. On the pilot this caught a 6.3 GB copy of the entire tender zipped inside itself. Corrupt or encrypted files are itemised in the intake report — never silently skipped.
- **Triage.** File-type census; per-PDF page census with text-layer detection (vector text, vector without text, or raster scan per page range); junk filtering (backup files, plot logs, thumbnails).
- **Identify.** Sheet identity is read from the **title block** — drawing number, title, revision, date, scale, discipline — never from the filename, because folder names and filenames routinely lie. Multi-sheet PDFs are split on detected title blocks, not page counts.
- **Resolve.** Where several instances of the same sheet number exist (and they always do), the **Effective-Set Resolver** elects the governing revision using addendum references, revision tokens, title-block dates and revision-history tables — producing one authoritative set, with every superseded instance kept on record for the audit trail and later delta bills.

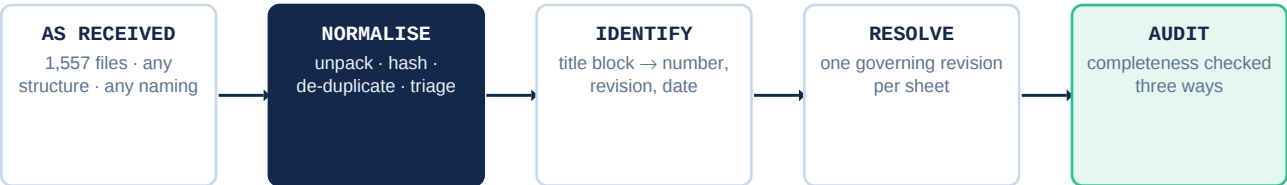


Figure 3.1 – Stage 01 pipeline: from raw dump to governed, audited set.

Completeness, checked three ways

A completeness check is only as good as its definition of “complete”. The engine uses three independent sources, in decreasing order of authority — and the third is what makes it unusual: it can find documents **the consultant never admitted existed**.

Source of truth	What it catches
1 · The project’s own registers — the “list of drawings” sheets in every discipline set and the volume indexes.	Sheets the register promises but the delivery lacks.
2 · Cross-reference closure — every callout on every sheet (“SEE DETAIL 5/A-501”, “REFER S-06”) must dereference to a received sheet.	Dangling references — details and sections cited but never issued, even when no register lists them.
3 · The required-set checklist — a curated ontology of what a project of this type needs to price each bill section (door schedules for openings, lintel schedules for blockwork, and so on).	Documents nothing references but pricing requires: pricing preambles, affection plans, geotechnical annexes. On the pilot this caught referenced-but-absent preambles PP/1–28.

How this stage learns: each new consultant's conventions (title-block geometry, numbering scheme, revision tokens) are confirmed once by a human and stored as a profile. The second tender from the same consultant intakes with no setup at all.

PART III · STAGE 01 · DELIVERABLES

Stage 01 — deliverables & operational benefit

Deliverable	What it contains	Why it matters
Intake report	Full file inventory with hashes; duplicates and junk itemised; per-file readability status; unreadable/encrypted items listed with reasons.	You know — provably — what you received. Disputes about “what was issued” end here.
Effective-set register	One governing revision per sheet with the evidence for each election; supersede graph; ambiguous cases queued for a 15-minute human confirmation.	Everyone measures from the same governing set. The classic “priced off the superseded drawing” failure becomes structurally impossible.
Completeness report	Findings from all three completeness sources, each with severity and the evidence trail; coverage matrices (levels × disciplines × view types).	Gaps surface on day one — while there is still time to demand the documents.
Document-request register	A consultant-ready, numbered request list for every missing item, with the citation that proves it is missing.	Requests carry evidence, not opinion — consultants respond faster to “cited by 4 sheets, not received” than to “we think something is missing”.

Operational benefit

- **The first week comes back.** The study that costs an estimating team its first 4–5 days is on the table the same day the tender arrives — as a report, not as tired recollection.
- **Risk is visible before money is spent.** Bid/no-bid (Stage 03) is briefed by a complete inventory of what is missing and ambiguous, not by a hopeful skim.
- **The record is court-grade from day zero.** Hashes, registers and supersede graphs are exactly the evidence that wins later arguments about what was issued and when (Stage 08).

MEASURED ON THE PILOT

1,557 files across five volumes, three addenda and three clarification rounds: unpacked, fingerprinted, identified and resolved to ~500 governing sheets — with zero unresolved elections and the full study report issued **same-day**.

PART III · STAGE 02

02. Red Flags & Clarifications

Every gap and contradiction found before the query deadline — evidenced, numbered, and written up as questions the consultant can actually answer.

What happens, in detail

With the governed set established, a catalogue of **~40 named validation checks** sweeps everything the engine has extracted — drawings, schedules, specifications, and the bill template itself. Each check is deterministic code with typed inputs: it either fires with evidence or it does not fire. Nothing depends on an AI “noticing” a problem. The catalogue spans eight families:

Check family	What it hunts	Example from live tenders
A · Completeness	Missing sheets, dangling references, absent schedules, spec/BOQ gaps	Per-board DB schematics cited by 4–5 sheets each — never issued
B · Revisions	Duplicate sheet numbers, ambiguous ordering, addendum references to un-received revisions	28 ambiguous duplicate copies resolved through one standing directive
C · Single-sheet	Dimension chains vs stated overalls, schedule-internal arithmetic, wrong project’s template	A finishing schedule headed with a different project’s plot number — template reuse caught by the project-identity check
D · Cross-sheet	Grid and level consistency between sheets, callout targets, tag↔schedule joins	Two level datums where sections disagree with each other
E · Cross-discipline	Architectural vs structural vs MEP geometry for the same rooms and elements	Pump-room dimensions differing between structural and electrical sets
F · Count reconciliation	Drawn tags vs schedule quantities per mark, symbol counts vs equipment schedules	927 drawn tags tallied; 57 genuine drawn-vs-scheduled conflicts confirmed
G · Units & geometry	Unit sanity per CAD file, grid arithmetic, annotated vs measured dimensions	A schedule CAD drawn in inches inside a metric set — the ×25.4 class, gated
H · BOQ / spec alignment	Bill sections without spec anchors, PC/PS items without references, duplicated bill pages	Pricing preambles referenced in the index, absent from the issue

Every check that fires produces a **finding**: a typed record with severity (blocker/major/minor/info), the observed values, and evidence pointers that dereference to the exact sheet locations. Findings then flow into two client-facing documents: a **technical query pack** (numbered questions with drafted text and figures cut from the actual drawings — every counted element ringed in place so the consultant can verify without opening a CAD file) and a **drawing-defect register** tabulating each conflict with its delta. When answers come back, they are logged against their findings and only the affected results re-run.

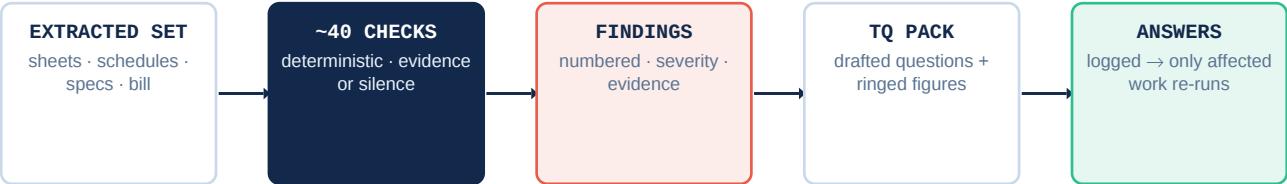


Figure 3.2 – Stage 02: from governed set to consultant-ready queries and back.

PART III · STAGE 02 · DELIVERABLES

Stage 02 — deliverables & operational benefit

Deliverable	What it contains	Why it matters
Findings register	Every issue as a numbered, severity-graded record with observed values, deltas and dereferencing evidence; lifecycle tracked from OPEN through RFI-SENT to RESOLVED.	Nothing lives in email threads or memory. The register is the single truth of what is wrong and what has been answered.
Technical query pack	Consultant-ready numbered queries with drafted question text and rendered evidence figures — counted elements ringed directly on drawing extracts.	Queries land before the deadline with proof attached. Consultants answer faster, and your professionalism is visible to the employer's team.
Drawing-defect register	Tabulated drawn-vs-scheduled conflicts and cross-discipline contradictions, each with both readings and the delta.	The complete, quantified picture of documentation quality — priceless at bid/no-bid and again in claims.
Provisional governing values	For every conflict, the value the bill will provisionally use, selected by a codified precedence ladder, flagged CONTESTED until the consultant answers.	Work continues while queries are outstanding — without hiding that the line is disputed.

Operational benefit

- **Queries become a weapon, not an afterthought.** On the pilot: 212 findings and 165 drafted queries, in before the deadline. The consultant's answers de-risk **your** price specifically — competitors who never asked keep pricing the ambiguity.
- **Reputation compounds.** Consultants and employers notice the bidder whose queries arrive early, numbered, and with ringed evidence. That perception carries into award decisions and into every future tender.
- **Nothing found is ever lost.** Every conflict caught at tender stage is already documented and dated for the claims file no one wants to need (Stage 08).

WHY SPEED MATTERS HERE

The query deadline is fixed. Finding a contradiction in week one instead of week three is the difference between a consultant's answer (their risk) and a pricing assumption (your risk). Stage 02 exists to move risk across that line, at scale.

PART III · STAGE 03

03. Bid / No-Bid Decision Room

The most expensive sentence in contracting is “we may as well finish the bid now”. Stage 03 makes the go/no-go call data-backed, auditable, and cheap enough to make properly on every tender.

What happens, in detail

- **Commercial-conditions extraction.** The conditions of contract are parsed into a risk scorecard: liquidated damages and caps, retention and release, payment cycle and certification lag, escalation provisions, bond and guarantee requirements, unusual risk transfers. Each entry carries its clause citation, so the review argues from the document, not from summaries.
- **Package-risk scoring.** Stages 01–02 quantify documentation risk: how much is missing, how many contradictions, what remains unknown. A tender with 200 open issues and an unmovable deadline is a different bet from a clean set — now that difference is a number, not a feeling.
- **History comparison.** The opportunity is compared against your own record: similar projects by type and scale, their outcomes, their realised (not tendered) margins, and what went wrong. This is the institutional memory that usually retires with senior staff, made queryable.
- **The recommendation memo.** One page: bid or no-bid, with conditions attached (price the LD exposure; bid only if TQ-012 is answered; load the bond cost) and every input appended. If the answer is no-bid, the reasons are on paper — reviewable when the market shifts.

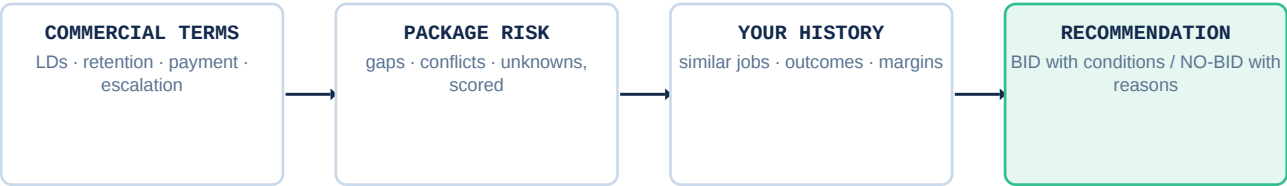


Figure 3.3 – Stage 03: three evidence streams, one defensible call.

How this stage learns: every processed tender enriches the history base — outcomes, margins, consultant behaviour, addendum frequency — so the scoring sharpens with each decision the desk makes.

PART III · STAGE 03 · DELIVERABLES

Stage 03 — deliverables & operational benefit

Deliverable	What it contains	Why it matters
Commercial risk scorecard	Every onerous condition extracted with clause citation: LDs, retention, payment terms, escalation, bonds, unusual transfers.	The board argues from clauses, not impressions. Nasty surprises surface before the bid, not in month three of the contract.
Package-risk summary	Documentation-quality metrics from Stages 01–02: missing documents, open conflicts, unknowns — trended against your other tenders.	A quantified read on how dangerous this particular tender’s paperwork is.
History comparison	Nearest comparable past projects with outcomes and realised margins.	Decisions informed by what actually happened last time — not what is remembered.
Recommendation memo	BID with explicit conditions, or NO-BID with explicit reasons; all inputs appended.	The call is auditable and consistent. When it is revisited later, everyone can see exactly what was known at the time.

Operational benefit

- **4–5 days becomes one.** The senior-time bottleneck at the front of every tender collapses, because the evidence assembly is already done by Stages 01–02.
- **You can afford to screen everything.** When evaluation is cheap, no winnable tender is declined unexamined — and no doomed one absorbs three weeks of desk time out of sunk-cost momentum.
- **Bid quality rises selectively.** The tenders you do enter get the full machine behind them, chosen because the evidence said they were winnable.

DISCIPLINE, ENCODED

Most contractors have a bid committee; few have bid-committee **records** that survive scrutiny two years later. Stage 03 turns the go/no-go meeting from a conversation into an artefact.

PART III · STAGE 04

04. The Bill of Quantities

The centre of the suite: a bill where every line shows its working. Incomplete is allowed — wrong is not.

What happens, in detail

The engine operates in two modes, matching the two commercial situations a contractor faces:

- **Verify mode** — the tender issued a BOQ (usually unpriced). The engine re-measures every line against the drawings, confirming or challenging the issued quantities before you price them. Where the issued bill is items-only, the engine quantifies the item universe independently — on the pilot it reproduced 46 of 47 opening items in a section purely from the drawings.
- **Generate mode** — no bill was issued (lump-sum risk is yours). The engine produces the bill from scratch: trade sections, item descriptions phrased through guarded templates, quantities from measured geometry and harvested schedules.

The anatomy of a bill line

Every quantity item the engine produces carries five mandatory attributes — this is the contract that makes the bill defensible:

Attribute	Example	What it guarantees
Quantity + unit	39,418 m²	The number itself — never typed by an AI, always computed.
Formula	$\Sigma(L \times H)$ – openings > 0.5 m²	A re-executable expression. Anyone — including the independent audit — can recompute the line.
Source	A-SEC-200 · REV 2 · grid C–E	The governing sheet and location the inputs came from, dereferenceable to the pixel or CAD entity.
Confidence class	SCHEDULE / MEASURED / DERIVED / INFERRED / MANUAL	An honest grade of how the number was obtained — printed in the bill, not hidden in metadata.
Status flags	CONTESTED → TQ-014	Lines resting on an unresolved conflict are visibly linked to the live query until the consultant answers.

The confidence classes enforce a discipline most manual bills never state: **SCHEDULE** means the value was read from an issued schedule (highest authority under the precedence ladder); **MEASURED** means computed from CAD geometry; **DERIVED** means calculated from other measured values; **INFERRED** means an assumption was applied — and the assumption is itself a numbered, citable record in the assumptions ledger; **MANUAL** means the engine declined to guess and says why (unmeasurable source, missing input, pending answer). The bill is permitted to be incomplete; it is never permitted to be silently wrong.

Output formatting is a separate, learned layer: a **template learner** ingests one of your existing workbooks and reproduces your sections, numbering, preamble conventions and summary structure — so the deliverable drops into your existing pricing workflow without retraining anyone.

PART III · STAGE 04 · DELIVERABLES

Stage 04 — deliverables & operational benefit

Deliverable	What it contains	Why it matters
The BOQ workbook	All trade sections in your own format, every line carrying quantity, formula, source, confidence and flags; contested and inferred lines visually marked.	Drops straight into pricing. Reviewers see instantly which lines deserve attention.
Review queue	The flagged subset — CONTESTED, INFERRED, MANUAL — with provenance links for one-click verification.	Your QS reviews ~3% of the bill with full evidence instead of spot-checking 100% with none.
Assumptions ledger	Every assumption as a numbered record: what was assumed, why, its source class, which lines cite it.	Assumptions become managed objects — reviewable, challengeable, and revisable in one place.
Basis of estimate	Rule-pack version, governing document set, open findings summary, measurement conventions used.	The document that makes the bill defensible in adjudication and precious in claims.

Operational benefit

- **Weeks become days.** The measurement bulk of tendering is machine work; the pilot produced 5,934 measured items across all fifteen sections, 979 machine-quantified in the flagship trades with the remainder honestly graded.
- **Contingency shrinks without margin shrinking.** Enumerated uncertainty needs less padding: when you know exactly which 3% of lines are soft, you stop pricing the other 97% as if they were.
- **Every review conversation gets shorter.** “Where did this number come from?” is answered by the line itself.

THE DESIGN RULE THAT GOVERNS THIS STAGE

Incomplete is allowed. Wrong is not. A bill that admits what it could not measure — with reasons — is commercially safer than a complete-looking bill hiding guesses. Every competitor demo shows a full bill; ask instead which lines *not* to trust, and watch.

PART III · STAGE 05

05. The Bill of Materials

The BOQ answers “what work”; the BOM answers “what to buy”. Both are computed from the same measured spine — nothing is ever measured twice, so they can never drift apart.

What happens, in detail

The BOM is not a spreadsheet derived by a different person on a different day — it is a branched execution layer consuming the exact quantity records the BOQ used, then diverging through its own materials logic:

- **Assembly resolution.** Each bill element class maps to an assembly recipe (a m² of 200 mm blockwork = blocks + mortar + reinforcement mesh where specified). An element with no recipe raises a finding — it is never silently skipped.
- **Material explosion.** Recipes explode through **consumption norms** — blocks per m², mortar per m², formwork ply per m² per reuse cycle. Norms are your own where you have them (purchase-department sheets outrank industry standards outrank assumptions), and every norm carries its source.
- **Policy application.** Wastage percentages, lap allowances, formwork reuse counts — each policy cites its source document; assumed policies are quarantined in the assumptions ledger and marked INFERRED in the output.
- **Procurement aggregation.** Net material quantities convert to purchase units — bags, pallets, standard lengths, tonnes by bar diameter — and aggregate by SKU across the whole bill, plus consumables and QC items (curing compound, tying wire, test cubes).
- **Two-way reconciliation.** Every BOM line traces to its contributing BOQ lines and every BOQ line to its material consequences, cross-footed in both directions. If the two documents ever disagree, the reconciliation names the exact line — before procurement does.

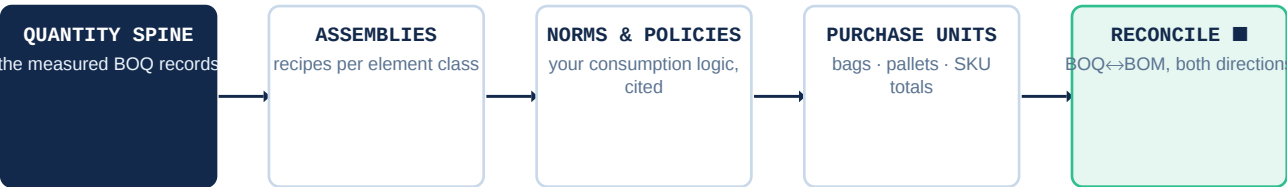


Figure 3.5 – Stage 05: one measured spine, exploded into what you buy.

How this stage learns: your norms library is signed off once per client and reused on every project; every correction a buyer makes becomes a norm revision the next tender inherits.

PART III · STAGE 05 · DELIVERABLES

Stage 05 — deliverables & operational benefit

Deliverable	What it contains	Why it matters
The BOM workbook	Materials by trade and by SKU, in purchase units, with per-line provenance back through assembly and norm to the originating bill lines.	Procurement plans from a document it can trust — and challenge, because every figure shows its derivation.
Norms & policies register	The full consumption-norm library used, each entry with source authority (purchase dept > client SOP > industry standard > assumed).	Your materials logic becomes an owned, versioned asset instead of tribal knowledge in one buyer's head.
Reconciliation report	BOQ↔BOM cross-footing both directions, with any orphan or mismatch named.	The QS desk and the buying desk stop drifting apart — structurally.
Consumables schedule	The easily-forgotten items (curing compound, spacers, tying wire, test cubes) computed from governing quantities.	The 2% of items that cause 20% of site purchase-order chaos, handled up front.

Operational benefit

- **Procurement starts the day the bill lands** — not after a second measurement exercise.
- **Buying quantities stop being disputed internally.** One spine means one truth; the reconciliation report is the referee nobody has to argue with.
- **Wastage and policy become visible costs.** When the +5% is a cited policy rather than a habit, management can actually decide it.

PART III · STAGE 06

06. Addendum & Revision Watch

The single highest-leverage feature in the tender workflow: a late addendum becomes an hour-long diff instead of a week of panic re-measurement.

What happens, in detail

- **Re-resolution.** A new addendum or revised drawing enters exactly like original intake: identified from its title block, slotted into the supersede graph, and the effective set re-elected. The engine also parses the addendum document itself — which sheets it claims to add or replace — and verifies those claims against what actually arrived.
- **Dependency-tracked invalidation.** Because every extraction and quantity cites its source sheets, the engine knows precisely which results the change touches. Only those recompute — on the pilot, a full warm re-run of all fifteen sections takes minutes, so the incremental case is near-instant.
- **The delta bill.** The output is not a new 6,000-line bill to re-review, but a difference report: items added, items removed, quantities changed (from → to), findings opened or closed, and the materials impact — the delta BOM follows automatically.
- **Reference verification.** The addendum's own citations are checked: if it refers to an updated mechanical layout that was never attached, that absence is a finding within minutes of receipt — while there is still time to demand the document.

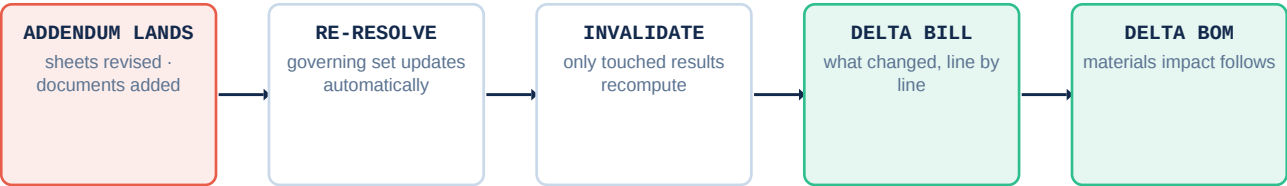


Figure 3.6 – Stage 06: the revision loop. The same machinery serves construction-phase revisions after award.

PART III · STAGE 06 · DELIVERABLES

Stage 06 — deliverables & operational benefit

Deliverable	What it contains	Why it matters
Delta BOQ	Per-line changes with before/after quantities, new and deleted items, and the finding or addendum reference that caused each change.	Re-pricing effort goes exactly where change happened — nowhere else.
Delta BOM	The materials consequence of every bill change, in purchase units.	Procurement adjusts orders from a one-page diff instead of a re-issued book.
Addendum verification note	What the addendum claimed vs what was received; missing attachments flagged as document requests.	You respond to the addendum's gaps while the consultant can still fix them.
Updated registers	Effective set, findings and query registers refreshed; superseded lines retained for the record.	The audit trail survives every revision — which is exactly when audit trails usually die.

Operational benefit

- **About an hour to a re-priced bill** after a late addendum — measured against the days of manual re-measurement it replaces, at exactly the moment the tender clock is loudest.
- **Final-48-hours agility becomes a bidding weapon.** You can absorb a week-three addendum calmly while competitors triage what to re-measure by hand.
- **The same machinery keeps working after award** — construction-phase revisions produce the same delta bills, which is where variation pricing begins (Stage 08).

WHY THIS IS ARCHITECTURALLY CHEAP

Revision watch is not a feature bolted on — it falls out of the evidence spine. Because every result cites its sources, “what does this change touch?” is a database query, not an investigation.

PART III · STAGE 07

07. Procurement & Subcontracting

From master bill to like-for-like vendor comparison — with every exclusion exposed before award, not at the first invoice.

What happens, in detail

- **Package splitting.** The master bill splits into trade-package BOQs — blockwork, MEP, finishes, or whatever package structure you buy in — each a complete, priced-format enquiry document.
- **RFQ generation.** Per package, an enquiry pack: scope description, the package bill, key specification extracts and commercial terms — issued to your vendor list.
- **Quote normalisation.** Returned quotes never match: different units, different exclusions, different scope interpretations. The engine aligns units, detects exclusions and scope gaps against the package bill, and prices the gaps so comparison is genuinely like-for-like.
- **The levelled matrix.** Vendors ranked on levelled totals — quoted price plus the priced cost of what each quote leaves out — with every adjustment traceable.
- **Rate library capture.** Every quoted rate lands in the rate library with its context (project, package, vendor, date), feeding future estimates and Stage 03 history.

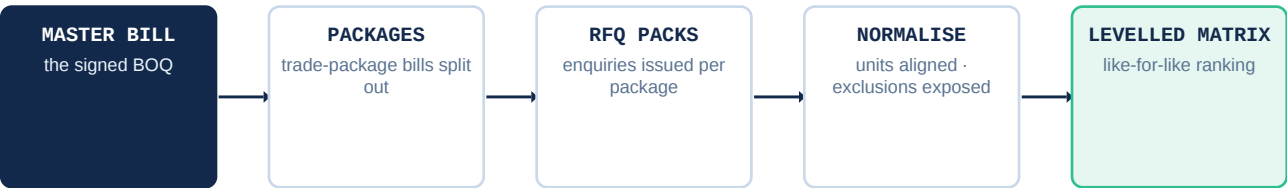


Figure 3.7 – Stage 07: the enquiry cycle, compressed and made comparable.

How this stage learns: the rate library compounds fastest of all the assets — every tender adds dozens of dated, contextual rates that sharpen the next estimate before a single enquiry goes out.

PART III · STAGE 07 · DELIVERABLES

Stage 07 — deliverables & operational benefit

Deliverable	What it contains	Why it matters
Trade-package bills	Complete per-package BOQs carved from the master bill, in enquiry-ready format.	Subcontract enquiries go out days earlier, so quotes return before pricing closes.
RFQ packs	Scope, bill, spec extracts and terms per package.	Vendors quote against defined scope — fewer qualification games later.
Levelled comparison matrix	As-quoted vs levelled totals per vendor, with every exclusion priced and every adjustment cited.	The buying decision is defensible to management and auditors — best like-for-like, not loudest discount.
Rate library entries	Every quoted rate captured with context and date.	Institutional pricing memory that compounds with every tender.

Operational benefit

- **The RFQ cycle compresses from weeks to days**, because package bills exist the day the BOQ is signed.
- **Exclusions surface before award** — the classic “that was excluded, see note 7” invoice surprise is caught in the matrix instead.
- **Every negotiation starts better informed**, with the rate library’s history behind it.

PART III · STAGE 08

08. Award & Beyond

The evidence spine outlives the tender. Everything documented to win the job keeps working after you have won it — the same analysis, used twice.

What happens, in detail

- **Variation pricing from delta bills.** Construction-phase revisions flow through the same Stage 06 machinery: each revised drawing produces a delta bill that is, in substance, the quantum of a variation — computed the day the revision lands, not weeks later.
- **Claims and EOT files from the register.** Every discrepancy, gap and contradiction found at tender stage is already a numbered, dated, evidenced record. When a dispute arises, the file assembles from the registers instead of being reconstructed from memory and old emails.
- **Remeasurement to final account.** For remeasured contracts, the same measurement engine runs against as-built revisions, with the same provenance discipline, to final account.
- **The institutional flywheel.** The project’s profiles, norms, rates and outcomes fold back into the library — Stage 03 gets a new comparable, Stage 07 gets new rates, and the next tender starts further ahead.

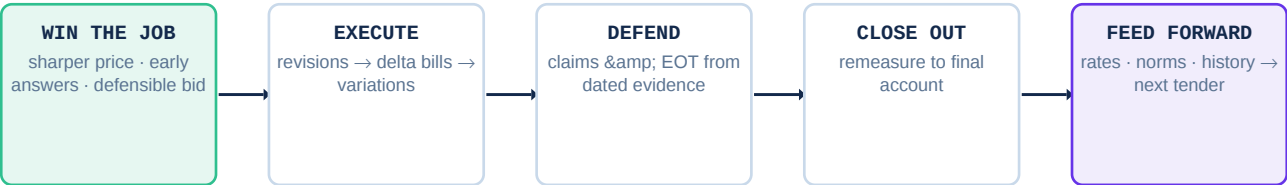


Figure 3.8 – Stage 08: from award through final account, on the same spine.

Deliverables & operational benefit

Deliverable	Why it matters
Variation quantum reports — delta bills per construction revision, with materials impact.	Variations priced while the record is fresh; entitlement supported by computation, not assertion.
Claims evidence packs — the relevant findings, queries, answers and dates, assembled per dispute.	The party with the contemporaneous record usually wins. You are always that party.
Remeasurement bills — progressive and final-account measurement on the same engine.	Interim applications and the final account argue from the same provenance the tender did.
Portfolio feedback — outcomes, rates and profiles folded into the libraries.	The self-improving loop closes: every completed project makes every future tender sharper.

THE ECONOMIC POINT

The analysis was paid for once, at tender stage. Stage 08 is where it pays back the second time — in variations recovered, claims defended, and final accounts closed without archaeology.

PART IV · TECHNICAL ARCHITECTURE

4. How the engine actually works

This part is written for the reader who wants to know why the numbers can be trusted — the pipeline, the data model, the role AI is permitted to play, and the five structural layers that stand between a language model and your bill.

4.1 The whole route on one board

The engine's complete processing flow fits on a single A3 board, reproduced below and issued alongside this document. Ovals convert files into readable form, boxes do work, diamonds are the reference material work is checked against, and the dashed line is the revision loop from Stage 06.

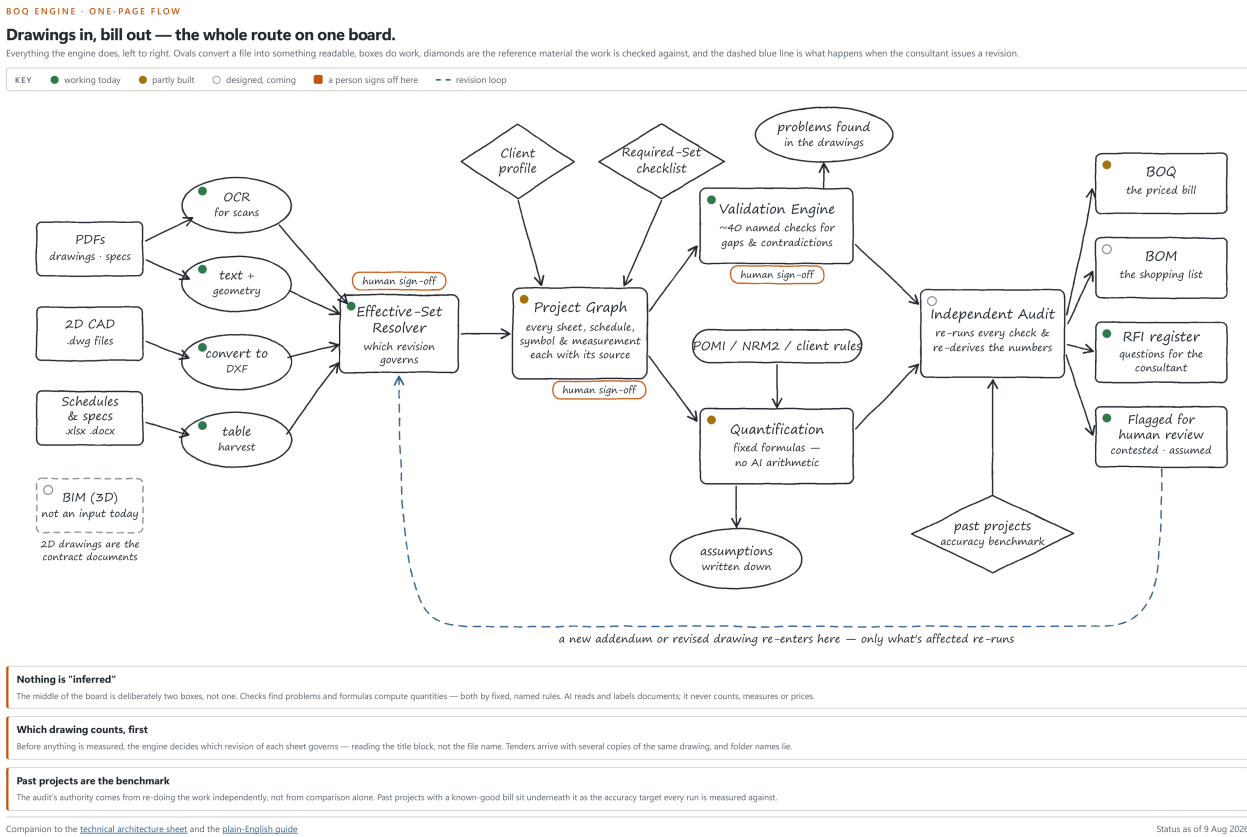


Figure 4.1 – the one-board engine flow (companion A3 sheet: BOQ-Engine-Flow-Diagram).

Reading it left to right: **inputs** (PDF drawings and specs, 2D CAD, Excel/Word schedules — with scans routed through OCR for text only, and BIM noted as a future input, since 2D drawings remain the contract documents in this market) flow into **conversion**; the **Effective-Set Resolver** decides which revision of each sheet governs; **semantic extraction** builds the project graph; the **validation engine** (~40 named checks) and the **quantification layer** (fixed formulas) do the analytical work; an **independent audit** re-runs and re-derives everything; and the deliverables issue as the BOQ, the BOM, and the query register. Two reference stores sit beside the route: the client profile and the required-set checklist; past projects sit beneath it as the accuracy benchmark.

PART IV · PIPELINE STAGES

4.2 The pipeline, stage by stage

Stage	What it does	Key guarantee
S0 · Intake	Recursive unpack, SHA-256 fingerprinting, duplicate & junk filtering, file census with text-layer detection.	Nothing is silently skipped; unreadables are itemised.
S1 · Conversion	CAD converted locally to measurable vector form; multi-sheet PDFs split on title blocks; scans routed to OCR for text only.	Geometry is never measured from pixels — raster sheets trigger a document request, not a guess.
S2 · Registry	Title-block identity, register parsing, addendum parsing, supersede graph, election of the governing instance per sheet.	Identity from title blocks, never filenames; every election carries its evidence.
S3 · Extraction	Grids & levels, layer semantics, legends, schedules, notes, room polygons, geometry measurements, dimension harvest — into the project graph.	Every extracted fact carries provenance and a confidence grade.
S4 · Completeness	The three-source completeness engine (register, cross-reference closure, required-set ontology).	Finds what the consultant never admitted was missing.
S5 · Consistency	The ~40-check catalogue across eight families (single-sheet to cross-discipline to bill/spec alignment).	A finding may only be produced by a named, versioned check with a deterministic evaluator.
S6 · Rule packs	POMI/NRM2 base + client amendments + numeric conventions compiled into a frozen, hash-stamped pack.	Every bill states which rule-pack version priced it.
S7 · Quantification	Pure per-trade functions consuming graph + rule pack, emitting quantity items with formula strings.	No AI computes anything. Contested inputs propagate their flags.
S8 · BOQ + BOM	Twin branches from one quantity spine: bill assembly with the template learner; BOM explosion through assemblies/norms/policies.	Two documents, one truth, reconciled both directions.
S9 · Audit	Independent re-execution of every check, evidence dereferencing, re-derivation of material lines, rule-compliance and format checks.	The audit is itself tested with deliberately corrupted canary bills.
S10 · Watch	Revision intake, dependency-tracked invalidation, delta BOQ and delta BOM.	Weeks of revision churn become an hour-long diff.

PART IV · DATA MODEL

4.3 The project graph and the provenance contract

Everything the engine knows about a tender lives in one queryable store — the **project graph**: sheets and their instances, grids and levels, spaces, legend symbols, schedule rows, measurements, findings, assumptions, quantity items, bill lines, material items, audit verdicts, human-gate records and queries. Two disciplines apply to every row:

- **Provenance** — file, sheet, revision, locator (page region or CAD entity handle), and method. Any value can be dereferenced back to the exact place it came from, mechanically.
- **Confidence** — SCHEDULE / MEASURED / DERIVED / INFERRED / MANUAL, plus a CONTESTED flag linking disputed values to their live findings.

Alongside the graph sit the assets that make the engine **client-agnostic and self-improving**: the **project profile** (one per consultant: title-block geometry, numbering scheme, revision tokens, layer dictionary — confirmed once by a human, reused on every subsequent tender) and the **required-set ontology** (per project type, what documents each bill section needs — the codified version of a senior QS's completeness instinct). New consultant = new profile, engine code untouched; profiles accumulate into a library and the marginal setup cost falls with every project.

4.4 Where documents disagree: the precedence ladder

Real tenders contradict themselves, and a bill must still be produced while queries are outstanding. Every conflict therefore carries a **provisional governing value** chosen by a codified ladder — and the affected lines stay flagged CONTESTED until the consultant answers:

Rank	Governs over
1 · Addendum / clarification response	everything below
2 · Later revision of the same sheet	earlier revisions
3 · Stated (annotated) dimension	scaled or measured geometry
4 · Schedule row	plan-derived tag readings
5 · Larger-scale detail	smaller-scale plan
6 · Specification text	drawing general notes
7 · CAD vector geometry	vectorised-PDF geometry

The ladder mirrors standard contract order-of-precedence and is configurable per contract. Its purpose is honesty under pressure: the engine never silently picks a value — it picks, states the rule, and flags the line.

PART IV · THE TRUST BOUNDARY

4.5 Exactly where AI is allowed — and where it is not

Large language models are exceptional readers and terrible accountants. The architecture assigns them four sanctioned roles — all guarded — and structurally excludes them from everything numeric:

Sanctioned AI role	The guard on it
Classification — sheet types, CAD layer meanings, symbol semantics.	Schema-validated output; human-confirmed once per consultant through the layer/legend gates.
Normalisation — turning a messy schedule grid into typed rows.	The verbatim-numeric guard : every number in the output must exist in the input, or the result is rejected and retried; on second failure it hard-fails to a human. An AI cannot “helpfully correct” a value.
Description phrasing — bill-line wording from structured attributes.	Slot-filling templates; the model may not introduce numbers at all.
Narrative — query letters and finding summaries.	Generated <i>from</i> structured records, never the reverse; the record is the truth, the prose is presentation.

An AI may also *propose* an anomaly during a final sweep — but a proposal only becomes a finding if it can be converted into a structured record whose evidence dereferences and whose values re-verify mechanically. Otherwise it lands in a low-trust observations annex, visibly separate from the findings register.

4.6 Five structural layers against hallucination

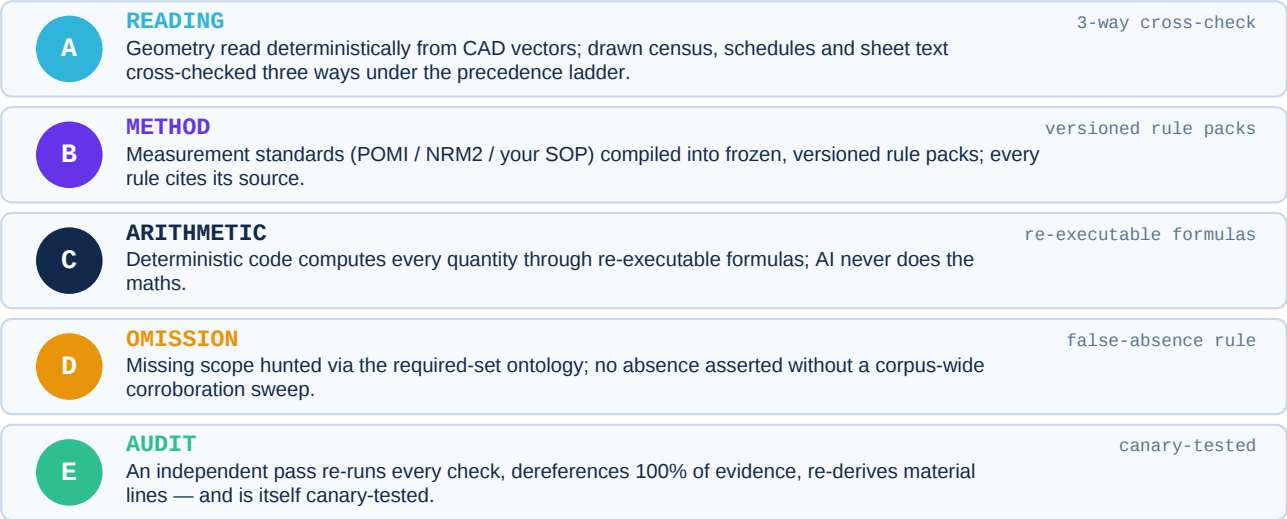


Figure 4.2 – the five defence layers. Each closes a distinct route by which a wrong number could reach the bill.

PART IV · THE INDEPENDENT AUDIT

4.7 Auditing the engine — and auditing the auditor

Before any deliverable issues, an independent audit process — separate from the generator, reading only the persisted artefacts and the raw sources — executes a fixed protocol:

- **Mechanical re-execution.** The entire check suite re-runs from the project graph; result sets must match exactly. Any difference indicts the pipeline itself and blocks issue.
- **Evidence dereferencing.** 100% of findings' evidence pointers are followed back to source bytes; the observed values must match what the source actually says.
- **Independent re-derivation.** Every bill line above materiality (and 100% of contested/inferred lines, plus a random sample of the rest) is recomputed from the graph — inputs re-queried, not trusted from the record. BOM lines re-derive from the quantity spine; reconciliations re-sum.
- **Rule and format compliance.** Formulas are checked against the cited rule pack (deductions, units, rounding); rendered workbooks are checked against the learned template, totals re-summed.
- **The canary discipline.** Periodically, a deliberately corrupted shadow bill is planted. An audit run that fails to catch its canary is itself invalid — the auditor is tested, not presumed.

THE RESULT ON RECORD

On the pilot engagement, independent re-derivation across all sampled sections returned **zero fabricated numbers** — and benchmark trades matched a client's own issued bill within $\pm 3\%$ (concrete & formwork) and $\pm 5\%$ (rebar).

4.8 Accuracy as an institution: golden corpora

Accuracy is not a launch claim; it is a regression suite. Every tender that comes with a known-good human bill becomes a **golden corpus**: a permanent benchmark the engine must keep matching within tolerance ($\pm 3\%$ concrete/formwork, $\pm 5\%$ rebar at current calibration) on every future software change. Hundreds of automated tests run against these corpora, and every new client project with a historical bill joins the set before the engine prices live work for that client. This is the mechanism behind the self-improving claim: the system does not just process your tenders — it is permanently examined against them.

PART IV · RULE PACKS

4.9 How measurement standards become code

Ask three estimators how to deduct openings from blockwork and you may get three defensible answers. The engine refuses to hold such answers as habits: every measurement convention is an explicit, versioned rule inside a **rule pack**, and every bill states which pack priced it.

What a rule pack contains

Layer	Contents	Example entries
Standard base	POMI or NRM2 (others as packs) — the published measurement rules encoded with their clause references.	Opening-deduction thresholds; measurement units per element class; what is deemed included in each item.
Client amendments	Your in-house SOP layered over the base — the deviations every contractor accumulates.	“Measure blockwork gross below 1 m² openings”; lap allowances; company rounding conventions.
Numeric conventions	The quiet parameters that move money.	Wastage policy by trade; formwork reuse counts; working-space allowances; rebar lap percentages.

Every rule is a typed record: {id · source: “POMI \$X” | “CLIENT-SOP \$Y” | “ASSUMED” · parameters · applies_to}. Three consequences follow. First, the independent audit can check **rule compliance**, not just arithmetic — a formula that deducts at the wrong threshold is caught even though its maths is perfect. Second, alternate standards are alternate packs on the same engine: switching a market or a client from POMI to NRM2 is configuration, not development. Third, disputes about method become citations — “priced per rule R-114, source CLIENT-SOP §3.2, pack v4” — which is a sentence that ends arguments.

WHY THIS MATTERS COMMERCIALLY

Most quantity disputes are really method disputes. A bill that cites its measurement rules line by line moves every such dispute from opinion to reference — in adjudication, in subcontract negotiation, and in claims.

4.10 The seven human gates

The suite is semi-autonomous by design — *autonomy is earned, never assumed*. Seven named gates put a human decision exactly where judgement belongs, and every gate action is minted as a permanent record: who confirmed what, and when.

Gate	The human confirms...	Frequency
G-REV	Governing-revision elections where the resolver found genuine ambiguity.	Short queue per tender; rare after profiles mature.
G-LAYER	The proposed CAD-layer semantic map (which layers mean walls, doors, tags...).	Once per consultant convention — ~15 minutes.
G-LEGEND	Symbol-meaning mappings harvested from drawing legends.	Once per consultant convention.

Gate	The human confirms...	Frequency
G-TRIAGE	Finding severities and which queries are dispatched to the consultant.	Per tender — the QS stays editor-in-chief of what leaves the building.
G-RULES	The rule-pack amendments and assumption-class rules.	Once per client; re-visited only on change.
G-TEMPLATE	The learned output format against a sample of your workbook.	Once per client format.
G-NORMS	The consumption-norms library behind the BOM.	Once per client; revisions versioned.

The gate ledger doubles as the management dashboard described in Part VI: because every human touchpoint is a recorded artefact, “where is tender X stuck?” has a factual answer at any moment.

PART V · SECURITY & DEPLOYMENT

5. Confidentiality by architecture

Tender documents are among the most commercially sensitive artefacts a contractor holds — your own rates and norms more sensitive still. The suite treats confidentiality as an architectural property, not a policy promise.

5.1 Local-first processing

The heavy machinery — CAD conversion, OCR, geometry measurement, validation, quantification — runs on controlled machines. No drawing is sent to a third-party service by default; any cloud-assisted step exists only behind a **written, per-project consent flag**. If consent is never given, the pipeline runs fully local and simply reports any conversion it could not perform.

5.2 Integrity and isolation

Every file is SHA-256 fingerprinted at intake, so any later substitution or corruption is detected rather than assumed away — and the fingerprints double as the evidence base for “what was issued” disputes. Projects run in isolated workspaces: one client’s drawings, norms, rates and profiles are never visible to another client’s processing, and never used to improve anyone else’s results.

5.3 Data residency, your way

Deployment option	What it means	Best for
In-country cloud	Dedicated environment in a cloud region inside your jurisdiction, aligned with UAE, KSA or India data-residency law; encrypted at rest and in transit; access-controlled and audit-logged.	Fast start, managed operations, multi-office access.
Fully on-premises	The entire pipeline inside your network, on your hardware; nothing leaves the building; model inference local or via private in-region endpoints under your control.	Strict IP policies, government or defence-adjacent work, maximum control.

5.4 What the AI models actually see

Because the pipeline is deterministic-first, model calls are narrow by construction: a table to normalise, a layer name to classify, a finding to phrase. Models never receive the tender wholesale, and inference can be pinned in-region or on-premises where residency demands it. Combined with the verbatim-numeric guard (Part IV), this means a model can neither see your tender at large nor *write* a number into your bill.

THE ENGAGEMENT STANCE

Every engagement — including the pilot — begins with an NDA before a single file moves. Deployment choice, retention terms and deletion-on-request are contractual, and the residency claim is designed to align with local data-protection law rather than argued after the fact.

PART VI · OPERATIONAL IMPACT

6. Where the weeks go — and come back

Machine hours replace estimator-weeks at every stage; the desk keeps the judgement. The figures below are pilot-observed; a calibration run on your own tender sets your numbers.

6.1 Stage-by-stage time impact

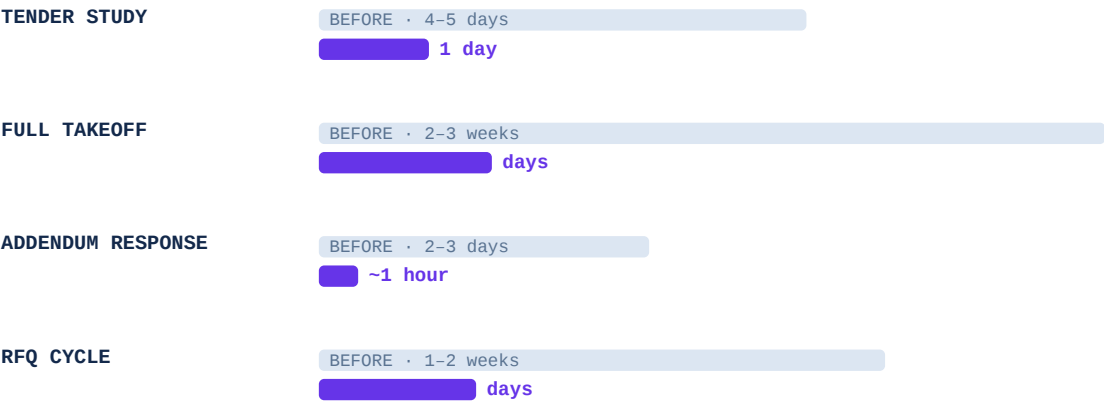


Figure 6.1 – before/after per stage (bar lengths to scale of elapsed time).

6.2 The cumulative picture

Summed across a typical tender, the elapsed analytical work compresses from roughly **22–33 working days to 5–7** — a cycle-time reduction of about **70–75%** — while the number of items actually examined goes *up*, because the machine checks everything rather than sampling. Two consequences follow: the same team can responsibly attempt two to three times the tender volume, and each tender gets deeper scrutiny than manual practice ever afforded it.



6.3 Team activity tracking

Because every human action in the pipeline is a recorded gate — layer confirmations, finding triage, assumption approvals, bill sign-off — management gains, as a side effect, a live view of where every tender stands and where hours are actually going: which stages are waiting on the consultant, which on internal review, and which closed same-day. No timesheets, no chasing; the audit trail doubles as the progress dashboard.

PART VI · EXPECTED RESULTS

7. Expected results — and the convergence effect

Three commitments define a successful deployment, and all three are measurable rather than impressionistic.

Expected result	Target	How it is measured
Total time reduction	Analytical tender cycle compressed by ≈70–75%; addendum response to ~1 hour; study to 1 day.	Stage timestamps in the pipeline itself — the system clocks its own turnaround on every tender.
Post-calibration accuracy	±2–3% on benchmark trades against issued human bills, after the calibration tenders.	The golden-corpus regression suite: every tender with a known-good bill becomes a permanent benchmark the engine must keep matching.
Bounded manual review	Human review confined to the flagged ~3% of lines (contested / inferred / manual) — shrinking further as profiles, rule packs and norms mature.	The review queue is a counted artefact per tender; its share is tracked release over release.

7.1 Why the numbers improve with every tender

The manual-review share starts near 3% because first tenders spend their flags on unfamiliarity: new consultant conventions, unconfirmed layer maps, unsourced norms. Each of those, once confirmed, is confirmed forever. Profiles absorb the consultant’s quirks; rule packs absorb your measurement amendments; the norms library absorbs your materials logic; the rate library absorbs your market. The flags that remain are the ones that *should* remain — genuine document conflicts awaiting a consultant’s answer.

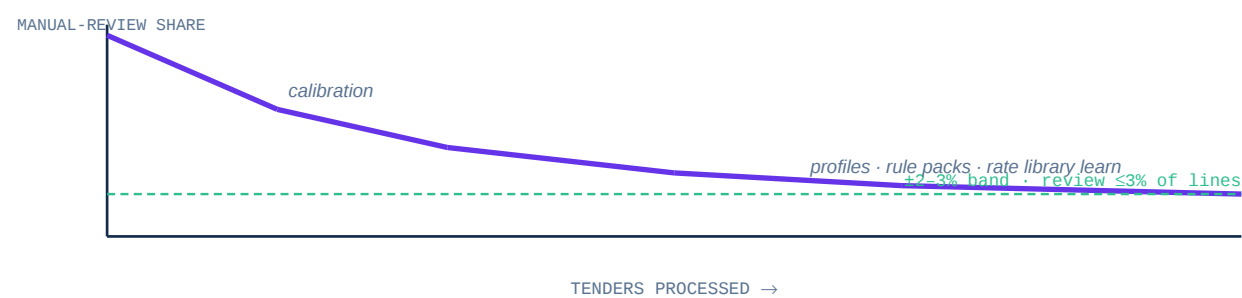


Figure 7.1 — the convergence effect: manual-review share falls toward the ±2-3% accuracy band as the system adapts to your business.

SELF-IMPROVING, STATED PRECISELY

The system does not vaguely “learn”. Four named assets compound — consultant profiles, rule packs, norms & rates, and the golden-corpus benchmark — and each is a versioned, inspectable artefact you own. Improvement is an audit trail, not a mystique.

PART VI · COMPETITIVE ADVANTAGE

8. Five ways this outbids the field

Speed is the visible half of the advantage. The deeper edge is informational: you bid knowing things competitors do not know, with evidence they do not have.

Mechanism	How it works	Who feels it
1 · More tenders, same team	Estimating capacity stops deciding which tenders you can attempt; screening is cheap and takeoff is machine work.	The MD — pipeline widens without headcount.
2 · Queries in first	Stage 02 lodges evidenced queries before the deadline; the consultant’s answers de-risk your price while competitors price the ambiguity.	The estimator — and the consultant, who now knows who the rigorous bidder is.
3 · Sharper price, same margin	Enumerated risk needs less contingency padding; you cut price where certainty is high and hold margin where it is not.	Adjudication night — the discount that used to be a gamble becomes a decision.
4 · Late-addendum agility	A week-three addendum is an hour-long diff; you re-price calmly in the final 48 hours.	The whole desk, at the worst moment of every tender.
5 · Claims-ready from day one	Every discrepancy is documented, dated and evidenced at tender stage — the record that wins variations and defends EOTs.	Commercial & contracts, for the life of the project.

And it compounds. Profiles, rule packs, norms, rates and history make every next bid faster and smarter than the last — an advantage that widens with use and cannot be bought off the shelf, because it is trained on your business, owned by you, and invisible to your competitors.

PART VII · THE TEST PILOT

9. Prove it on your own project — NDA first

Nothing in this document asks to be believed. The engagement is designed so every claim is tested against your own drawings, your own bills and your own clock before any commercial commitment.

9.1 The pilot, step by step

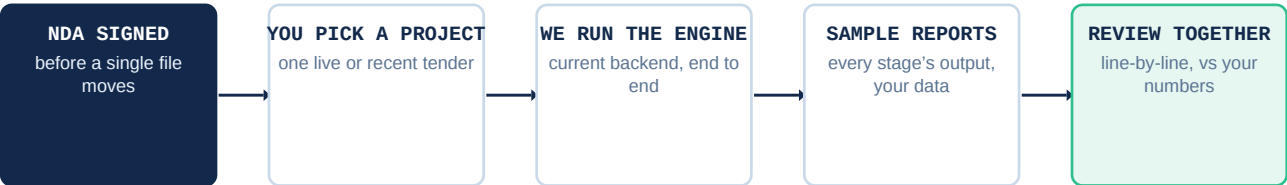


Figure 9.1 – the pilot path: evidence before commitment.

What you receive from the pilot

Sample deliverable	From stage
Intake & completeness study of your tender, with the document-request register	01
Findings register + technical query pack with ringed evidence figures	02
Commercial risk scorecard + bid/no-bid style recommendation memo	03
BOQ sections in your workbook format — quantities with formula, source and confidence per line	04
BOM extract, reconciled two-way against the sample bill sections	05
A delta bill from one of the tender’s actual addenda	06
A trade-package split + levelled quote comparison template	07
The audit certificate for the sample: what was re-derived and what matched	09/S9

9.2 What the pilot is for

- **Clarity** — you see every stage’s real output on your own project, not a demo dataset.
- **Confidence** — where a parallel manual takeoff exists, the two are compared line by line; divergences are walked back to evidence, in the room.
- **Calibration** — the pilot doubles as the first training pass: your formats, conventions and norms begin their profile, so a subsequent deployment starts warm.

THE ASK, PRECISELY

One NDA. One project’s drawings. Zero workflow change on your side — your team keeps working exactly as it does today while the engine runs in parallel. The review meeting decides everything that happens next.

ANNEX A · DELIVERABLE ANATOMY

Annex A — what the deliverables actually look like

Screenshots date; structures endure. This annex documents the exact anatomy of the three most load-bearing artefacts, so your reviewers know what to expect before the pilot pack arrives.

A.1 A finding, dissected

Every issue the engine raises is a typed record. Here is a real-shaped example (values illustrative), field by field:

Field	Example value	What it is for
ID / check	F-0042 · check E2 (cross-discipline geometry)	Traceable to the named, versioned check that produced it — findings cannot appear from nowhere.
Severity · status	MAJOR · RFI-SENT	Lifecycle tracked from OPEN through TRIAGED, RFI-SENT, ANSWERED, RESOLVED, VERIFIED.
Title	Pump-room size conflict: structural vs electrical	Human-readable summary for registers and meetings.
Observations	4950×4900 mm on S-201 R1 (CAD text) · 5500×6000 mm on EL-SOP-02 R0 (PDF text)	Both readings, each with sheet, revision, exact location and extraction method — dereferenceable evidence, not paraphrase.
Delta	550×1100 mm · 13.8% area	The size of the disagreement, quantified.
Governing value	4950×4900 · rule: schedule-over-plan + discipline owner	The provisional value the bill uses, and the precedence rule that chose it.
Affected quantities	QI-0311, QI-0312	The bill lines that inherit the CONTESTED flag until resolution.
Draft query	“Pump room is 4,950×4,900 on S-201 R1 but 5,500×6,000 on EL-SOP-02 R0. Please confirm governing dimensions.”	Ready-to-send text — the QS edits, approves, dispatches.

ANNEX A · DELIVERABLE ANATOMY (CONT.)

A.2 A bill line, dissected

Column	Example	Notes
Item ref	H.12	Your numbering scheme, learned from your workbook.
Description	Blockwork, 200 mm solid, in walls; ...	Phrased through guarded templates from structured attributes — the model cannot introduce numbers.
Quantity · unit	39,418 · m²	Computed, never typed.
Formula	$\Sigma(L \times H)$ – openings > 0.5 m²	Re-executable; the audit recomputes it from the graph.
Source	A-SEC-200 · REV 2 (+ 13 sheets)	Governing sheets; every input dereferenceable.
Confidence	DERIVED	SCHEDULE / MEASURED / DERIVED / INFERRED / MANUAL.
Flags	CONTESTED → TQ-014	Live link to the query register until answered.
Rule citations	R-114 (POMI §—) · R-201 (CLIENT-SOP)	The measurement rules this line obeyed.

A.3 The audit certificate

Issued with every deliverable package: which checks re-ran (and that their results matched exactly), what share of evidence pointers were dereferenced (100% of findings), which bill lines were independently re-derived and their agreement, rule-compliance and format-compliance results, the canary status of the audit itself, and the residual-risk list — every open contested line and assumption, so the signing QS sees the complete honest picture on one page.

THE PILOT PACK, THEREFORE

When Part VII promises “sample output reports of each stage”, this annex is the contents list: study report, findings register, query pack, scorecard memo, bill sections in this anatomy, BOM extract, a delta bill, and the audit certificate over them all — generated from **your** drawings.

ANNEX B · ONBOARDING ROADMAP

Annex B — from NDA to production, week by week

Indicative timeline for the in-country cloud path; on-premises adds your infrastructure provisioning lead time. “Your effort” is deliberately itemised — the design goal is that adoption costs your team hours, not weeks.

Phase	Duration	What happens	Your effort
0 · NDA & scoping	Days	NDA executed; pilot project nominated; drawing dump transferred as-is.	One decision, one upload.
1 · Pilot run	1–2 weeks	The engine processes the tender end-to-end; sample pack (Annex A) generated for every stage.	None — your team works normally.
2 · Review	One meeting	Line-by-line walk of the pack against your numbers; divergences traced to evidence live.	The estimating leads in one room, half a day.
3 · Calibration	2–4 weeks	Profiles confirmed (G-LAYER/G-LEGEND), rule pack amendments signed (G-RULES), norms library signed (G-NORMS), template learned (G-TEMPLATE); one or two more tenders run in parallel to seed the golden corpus.	~15 minutes per gate, once; parallel review as desired.
4 · Production	Ongoing	Every tender that lands gets the eight stages; accuracy and review-share tracked against the Part VI targets; quarterly reviews against the golden corpus.	Triage and sign-off — the judgement that was always yours.

Exit is as designed-for as entry: your data, profiles, norms, rule packs and registers are exportable artefacts you own. The system earns its place tender by tender, or it goes — with your assets in hand.

PART VIII · FAQ

10. Frequently asked questions

Collected from real conversations with contractors, consultants and IT reviewers. Grouped by theme; cross-references point into the body of this document.

Accuracy & trust

Q1 · How do we know the AI hasn't invented a quantity?

Because it structurally cannot. AI models never perform measurement or arithmetic anywhere in the pipeline — every quantity is computed by deterministic code through a re-executable formula, and every AI text-reading step is wrapped in a verbatim guard that rejects any number not literally present in the source. On the pilot, independent re-derivation returned zero fabricated numbers. See Part IV.

Q2 · What accuracy should we actually expect?

Current calibration on benchmark trades: within ±3% (concrete & formwork) and ±5% (rebar) of issued human bills. The deployment target after your calibration tenders is ±2–3%, tracked permanently by the golden-corpus regression suite — accuracy is a monitored metric, not a launch claim.

Q3 · What happens when the engine cannot measure something?

It says so. The line is marked MANUAL with the reason (unmeasurable source, missing document, pending query) and lands in the review queue. The bill is permitted to be incomplete; it is never permitted to be silently wrong.

Q4 · The drawings contradict each other. Whose number goes in the bill?

A codified precedence ladder (addendum over revision over stated dimension over schedule over plan...) selects a provisional governing value; the line is flagged CONTESTED and linked to the drafted query until the consultant answers. Nothing is silently chosen. See §4.4.

Q5 · Who checks the machine?

Two layers. An independent audit process re-runs every check, dereferences all evidence and re-derives the material lines before anything issues — and is itself tested with deliberately corrupted “canary” bills it must catch. Above that, your QS signs every deliverable; the review queue concentrates their attention on the flagged ~3% of lines.

Q6 · Do you support POMI / NRM2 / our in-house measurement rules?

Yes — measurement conventions are compiled rule packs, not hard-coded behaviour. POMI and NRM2 ship as bases; your amendments and numeric conventions are layered on, versioned, and cited by every bill that uses them. Switching standards is a pack change, not a software project.

Data & security

Q7 · Where do our drawings actually go?

By default, nowhere. Processing is local-first; any cloud-assisted step requires written per-project consent. Deployment options are in-country cloud (UAE / KSA / India aligned) or fully on-premises. See Part V.

Q8 · Does our data train anyone else's system?

No. Projects run isolated; your profiles, norms and rates are yours, never shared across clients, and never used to improve another party's results. The self-improvement described in this document happens inside your instance, on your assets.

Q9 · What does the AI model see?

Slices — a schedule table to normalise, a layer name to classify, a finding to phrase. Never the tender wholesale. Inference can be pinned in-region or run on-premises where residency requires.

Q10 · What happens to pilot data afterwards?

Contractual deletion on request, with the NDA governing throughout. If you proceed, the pilot's calibration assets (profiles, norms) carry forward as your property.

People & workflow**Q11 · Is this replacing our estimators?**

No — it replaces their measuring, not their judgement. The QS confirms mappings, triages findings, approves assumptions, prices, and signs. What changes is the ratio: hours shift from counting to judgement calls, negotiation and strategy — and the same desk covers two to three times the tender volume.

Q12 · How much of our time does onboarding take?

Per new consultant convention: roughly a 15-minute confirmation of the proposed layer/legend mapping, once. Per client: a one-time sign-off of measurement amendments and consumption norms. Both are remembered permanently in profiles and packs.

Q13 · Our tenders are messy — nested archives, scans, inconsistent naming. Will it cope?

That is the designed-for case, not the exception. Intake unpacks recursively, fingerprints, de-duplicates and identifies sheets from title blocks, never filenames. Scans route to OCR for text only — geometry is never measured from pixels; a raster-only sheet triggers a document request for the CAD instead of a guess.

Q14 · What file formats do you accept? Do we need BIM?

PDF (vector or scanned), DWG/DXF, Excel and Word are the working set — matching what tenders actually contain. BIM/IFC is a roadmap input; it is not required, because in this market the 2D drawings remain the contract documents.

Q15 · Does it integrate with our pricing software and ERP?

The BOQ and BOM render into your own workbook formats via the template learner, which is the integration surface most pricing workflows (CCS/Candy-style, Excel-based adjudication) actually use. Direct API/export integrations are scoped per deployment.

Commercial & engagement**Q16 · What does it cost?**

Commercials follow the pilot, deliberately: the pilot review establishes measured value on your own tender first, then pricing is proposed against it. The pilot itself is scoped under NDA with no obligation to proceed.

Q17 · How long from agreement to live operation?

The pilot runs on one project immediately. A production deployment (in-country cloud) follows calibration in weeks, not quarters; on-premises timelines depend on your infrastructure provisioning.

Q18 · What if the engine's output causes a costly error?

The operating model is draft-with-provenance: every deliverable is signed by a qualified QS before commercial use. The engine makes the review radically better informed — flags, formulas, and evidence on every line — and the audit certificate records exactly what was verified. Liability allocation is defined

contractually per engagement.

Q19 · Can we run it on only part of the workflow — say, just tender study and queries?

Yes. The stages are modular; many clients start with Stages 01–02 (study + red flags) and 06 (revision watch), then extend to full takeoff as confidence builds. The evidence spine is shared, so later stages inherit everything earlier ones produced.

Technical

Q20 · What infrastructure does on-premises need?

Modest by engineering-software standards: a current multi-core workstation or small server with ample storage for tender corpora. CAD conversion and OCR run locally; model inference can run on local hardware or private endpoints. Exact sizing is provided in the deployment guide per scope.

Q21 · Mixed units, bilingual drawings, odd title blocks?

Units are checked per file by a dedicated gate (it caught an inches-drawn schedule inside a metric set — the ×25.4 class). Title-block layouts are per-consultant profile assets. OCR handles the common bilingual annotation patterns of the region; genuinely unreadable content is flagged, never guessed.

Q22 · How many tenders can run at once? What about very large ones?

Projects are isolated workspaces and parallelise across them; within one tender, the pilot scale — 1,557 files, ~500 governing sheets, 5,934 measured items — processed end-to-end with warm re-runs in minutes. Larger sets scale with hardware, and the incremental engine means day-to-day work is always delta-sized, not corpus-sized.

Adoption & operations

Q23 · How disruptive is adoption for the estimating team?

By design, minimal. The pilot runs entirely in parallel — nobody changes how they work. In production, the team’s visible workflow is receiving better inputs: a study report instead of a raw dump, a review queue instead of a blank bill. Training is measured in hours (the gates and the review queue), not weeks, because nobody has to learn to operate the pipeline itself.

Q24 · We run several offices. Can they share one deployment?

Yes — projects are isolated workspaces under one instance, with access control per project. Profiles, rule packs, norms and the rate library are shared assets, so an estimator in one office inherits the calibration earned by another. That is the point: the institution learns once.

Q25 · What if the consultant simply doesn’t answer our queries?

Then you bid the way you always did — except now with a precise, priced statement of what remains unresolved. Contested lines carry provisional governing values selected by the precedence ladder, the exposure is enumerable at adjudication, and the unanswered queries are on the record — which matters enormously if those ambiguities later become variations or claims.

Q26 · How do updates and new checks reach us? Could an update change our numbers?

Releases are versioned, and the golden-corpus regression suite is the update gate: a release that moves any benchmarked quantity beyond tolerance does not ship. Rule packs are versioned separately from software, so method never changes silently — a bill always states the pack that priced it.

Q27 · Does the suite price the bill as well as measure it?

The suite deliberately stops at quantities, evidence and vendor-quote leveling — pricing strategy and margin remain human decisions. What it contributes to pricing is the rate library (every leveled quote, dated and contextual) and the confidence map that tells you where padding is unnecessary. Your adjudication stays yours.

ANNEX C · GLOSSARY

Annex C — glossary of terms

Term	Meaning
BOQ (Bill of Quantities)	The priced-format schedule of works: items, descriptions, quantities, units — the document a tender is priced on.
BOM (Bill of Materials)	The purchasing view: materials in purchase units, exploded from the BOQ through assemblies and norms.
TQ / RFI	Technical query / request for information — formal questions to the consultant; answers amend the tender basis.
Addendum	A formal tender amendment issuing new or revised documents; supersedes earlier issues.
Effective set	The engine-resolved answer to “which revision of each sheet governs” — with evidence per election.
Supersede graph	The recorded lineage of every sheet instance: what replaced what, and why.
Project graph	The queryable store of everything extracted from a tender — sheets, schedules, measurements, findings, quantities — all with provenance.
Provenance	The evidence trail on every fact: file, sheet, revision, exact location, extraction method.
Confidence class	The honesty grade on every value: SCHEDULE, MEASURED, DERIVED, INFERRED, or MANUAL.
CONTESTED	The flag linking a bill line to an unresolved document conflict and its live query.
Finding	A typed, evidenced record of a gap or contradiction, produced only by a named check.
Check catalogue	The ~40 deterministic validations across eight families (completeness through bill/spec alignment).
Precedence ladder	The codified order deciding which document governs when two disagree.
Rule pack	A frozen, versioned compilation of measurement rules: POMI/NRM2 base + client amendments + numeric conventions.
POMI / NRM2	Published measurement standards (Principles of Measurement International; RICS New Rules of Measurement 2).
Assembly / consumption norm	The recipe converting work into materials (blocks per m², mortar per m²...), each with a cited source.
Delta bill	The difference report after a revision: items added/removed, quantities from→to, materials impact.
Quote levelling	Normalising vendor quotes to like-for-like: units aligned, exclusions priced, scope gaps exposed.
Rate library	The accumulated, dated record of every levelled vendor rate — institutional pricing memory.
Golden corpus	A tender with a known-good human bill, kept as a permanent accuracy benchmark the engine must keep matching.
Canary bill	A deliberately corrupted shadow bill planted to prove the independent audit actually catches errors.

Term	Meaning
Human gate	A named, recorded human decision point (layer maps, triage, rules, template, norms, sign-off).
Verbatim-numeric guard	The rule that any AI-read number must exist literally in the source, or the output is rejected.
Evidence spine	The through-line of the suite: every fact, in every stage, carries its provenance and confidence.

CLOSING

11. The proposition, restated

Manual takeoff spends your scarcest asset — senior estimating judgement — on work a machine now does better: exhaustive reading, exhaustive checking, exhaustive measurement, with evidence on every number. The Construction Takeoff Suite returns that judgement to where it wins tenders: pricing strategy, risk decisions, and negotiation.

- Eight stages, one evidence spine — from the raw tender dump to final account.
- AI reads and labels; deterministic code counts and computes; your QS signs.
- Five structural layers stand between any AI model and your bill.
- Accuracy is a regression suite, not a promise: $\pm 2\text{--}3\%$ post-calibration, review bounded at $\sim 3\%$ of lines and shrinking.
- Security is architectural: local-first, hashed, isolated, deployed in-country or on-premises.
- And the whole claim is testable on your own project, under NDA, before any commitment.



SOCIAL BUBBLES IT SERVICES L.L.C
Dubai, United Arab Emirates
AI + Creativity + ROI

Next step: nominate one tender. We will bring the NDA.