

Industrial paperwork in the Gulf

What it costs, why the last decade of software did not fix it, and how to find out what yours costs.

One Industry AI · Industrial AI for the Gulf back office · First Settlement, New Cairo, Cairo, Egypt · oneindustry.ai. Written for an operations director, a shared-services head or a finance lead in a Gulf industrial operation. Ungated, and no client is named in it.

1. The sentence this document exists to answer

Ask an operations director in the Gulf what their back office costs and they will tell you a headcount. Ask what the paperwork inside it costs and the question does not have an answer, because nobody has ever counted.

That is not negligence. It is a measurement problem with a specific shape: the cost is distributed across people who each spend a fraction of their week on it, it never appears as a line item, and the work is invisible precisely because it always gets done. Month-end closes. The permit gets signed. The submittal gets approved. The cost of all that closing and signing and approving is real, and it is carried in overtime, in delay, in rework and in the decisions that get made late because the number they needed was still being typed.

This document is about that cost: where it sits, what is actually known about its size, and how to find out what yours is in two to four weeks rather than in a procurement cycle.

2. Where the paper actually is

The industrial back office runs on a small number of document types, and they are the same in every operation of this kind. Not similar. The same.

Finance and procurement. Supplier invoices and the three-way match against purchase order and delivery note. Purchase orders. Payment certificates. Credit notes and debit notes.

Contracts. The contract itself, its variation orders, requests for information, and the correspondence that decides which of the three governs a dispute two years later.

Operations and HSE. Permits to work. Method statements. Job safety analyses. Inspection reports. Non-conformance reports. Toolbox talk records.

Warehouse and logistics. Delivery notes. Goods received notes. Material submittals awaiting approval. Customs declarations. Certificates of origin and mill certificates.

Knowledge. Procedures, standards and specifications, in a revision structure that everybody agrees on and nobody can navigate.

Shared services. HR letters, salary certificates, leave requests, visa and iqama paperwork, medical claims.

Two properties of that list matter more than its length.

The first is that it is bilingual, unevenly. The contract is Arabic because the client is a government entity. The invoice is English because the supplier is international. The HR letter is both, on the same page, because it has to satisfy a ministry and an employee. No single document is reliably in one language, which means no tool that handles one language reliably handles the work.

The second is that most of these documents exist to be checked against another document. An invoice is not information, it is a claim, and its value is entirely in whether it matches the purchase order and the delivery note. A submittal is a claim about a material against a specification. A permit is a claim about a condition against a procedure. This is why the work is so expensive and why it resists the

obvious kinds of automation: the task is not reading, it is reconciling.

3. What is actually known about the cost

Three kinds of figure get quoted in this market, and they are not equally trustworthy. Being explicit about which is which is the only way a document like this is worth reading.

3a. Figures from delivered work

These come from engagements the people in this practice have delivered, measured against the process that preceded them. They are our own records rather than audited third-party findings, and they are single engagements rather than averages.

WHAT MOVED	FROM	TO
Everyday-language recall over a procedure library	34%	89%
Quality review coverage on a call centre	~2% sampled	100%
Forecast error on a demand series	42%	19%
Time to find an answer across separate repositories	~12 minutes	under 90 seconds
Gulf dialect spread on a voice system, best to worst	19 points	5 points

What these have in common is that each is a coverage or a latency figure rather than a cost saving. That is deliberate, and it is the honest form of the claim: coverage and latency are measurable inside the system, and a cost saving requires assumptions about what the freed time was worth. Those assumptions belong to the operation, which is why the calculators on the website ask the reader for them rather than supplying them.

3b. Figures from the market, attributed

Published estimates put the failure rate of enterprise AI projects somewhere between 70% and 95%. The figure is widely reported, the range is wide because the definitions differ, and the reason is not in dispute: the failures cluster around automating a process that was already broken, or building a capable system for a problem nobody quantified first.

The vocabulary itself is the finding. An industrial back office runs on customs declarations, permits to work, requests for information, material submittals, variation orders, delivery notes and method statements, and a market that describes itself in terms of models and platforms does not reach any of them. This document names them, because naming the document is the first step in measuring what it costs.

3c. A round number with no denominator

A figure with no engagement behind it and no denominator under it is a marketing artefact rather than a measurement, whoever publishes it. The test is simple and it applies to every figure in section 3a as much as to any other: name the engagement, name the measure, and say what it was before. A number that cannot survive those three questions should not be quoted, including ours.

4. Why the last decade of software did not fix this

Most Gulf industrial operations of any size already own the software that was supposed to solve this. An ERP. A maintenance management system. A document management system. Often a workflow tool bought on top of those to make them talk to each other.

None of it was wrong to buy and none of it addressed the problem, for four reasons that compound.

The systems hold records, not documents. An ERP is excellent at storing a validated invoice and has no view at all on the scanned PDF that arrived by email, which is where the work is. The document management system holds the PDF and has no idea what is in it. The gap between the two is a person with a keyboard.

Workflow automation assumed the data was already structured. A rule engine can route an invoice once the invoice is a set of fields. Getting from a photograph of a delivery note to a set of fields is the part that was never automated, and it is the part that consumes the headcount.

Bilingual was treated as a localisation problem. The interface was translated; the documents were not. A system that reads English invoices and Arabic contracts by translating the Arabic first loses exactly the thing a contract is read for.

And the process was never examined. This is the largest of the four. Almost every one of these systems was implemented over the process as it was described, not the process as it ran, and the two differ in ways that matter: the undocumented step where somebody checks a figure against a spreadsheet, the approval that is actually granted by WhatsApp, the exception path that handles a third of the volume.

The consequence is a familiar shape. The software is in place, the reporting is better than it was, and month-end still waits for three people retyping what a supplier already sent.

5. The four shapes, and why sector proof transfers

Across a library of documented engagements, industrial back-office problems reduce to four shapes. The shape is what determines whether a system will work, not the industry.

A human reads it and types it in again. Invoices, contracts, scanned paper, submittals. The document arrives, someone extracts what matters, and a system receives it. The failure mode is volume: the work scales linearly with the paper and the headcount does not.

The answer exists and nobody can find it. Procedures, standards, specifications. The information is in the organisation and the retrieval is the problem. The failure mode is confident error: a new engineer uses the 2019 revision because nobody could tell her which was current.

Too much arrives to check it all, so you sample. Inspections, calls, feedback, camera hours. A percentage gets reviewed and the rest is assumed. The failure mode is the unsampled remainder, which is where the expensive surprises live.

The same judgment, made differently every time. Risk scoring, prioritisation, approvals. A decision that should be consistent depends on who is making it and how busy they are. The failure mode is defensibility: an auditor asks why two similar cases went different ways and nobody can say.

Why this matters for reading anybody's case studies, including ours. A telecom's accounts-payable problem and a contractor's accounts-payable problem are the same shape: a document arrives, a human reconciles it against two others, a system receives the result. The industry differs and the engineering does not. So the right question to ask a vendor is not "have you worked in my industry" but "have you solved this shape, and what moved". The first question has a comfortable answer and tells you nothing. The second one is checkable.

6. What to measure before buying anything

The single highest-return thing a Gulf industrial operation can do about its paperwork costs nothing and takes a fortnight of somebody's attention. Count.

Five measurements, in order of how much they change a decision.

- 1. Volume by document type, per month.** Not total documents. Invoices separately from delivery notes separately from submittals, because they are different problems with different economics. Count from the system, not from memory: remembered volumes are reliably wrong in both directions, usually low for the routine documents and high for the annoying ones.
- 2. Minutes per document, by hand, measured rather than estimated.** Sit with the person doing it and time twenty. Then time twenty of the awkward ones, because the average is not what costs you: the exception path is.
- 3. The share that is an exception.** What proportion of documents cannot be processed on the normal path. This is the number that decides whether automation pays, and it is almost never on anyone's dashboard.
- 4. Where the wait is.** Month-end does not wait for the typing, it waits for the last document. Identify which document type is last and why, because a system that speeds up the other four types changes nothing about the close.
- 5. What a mistake costs, by type.** A mis-keyed invoice quantity, a missed variation order, an expired permit found by an auditor. These are not the same number and the difference decides where accuracy has to be near-perfect and where good enough is genuinely good enough.

With those five, a business case writes itself and is defensible. Without them, any business case is an assertion with arithmetic attached.

7. What to ask a vendor

Six questions worth asking before commissioning this kind of work. They apply to this practice as much as to anyone, and the answers are in the case studies on the website.

- **"Name the documents you have processed."** Not the capability, the documents. A firm that has done industrial back-office work says "supplier invoices, three-way match, delivery notes, material submittals". A firm that has not says "unstructured documents".
- **"What was the exception rate, and what happened to the exceptions?"** Any system reaches a document it cannot resolve. The answer that matters is where it went and who looked at it, not the

accuracy figure on the documents it could.

- **"Show me a before and an after for one engagement, with the measure named."** One engagement, not an average. Coverage from 2% to 100% is a claim you can interrogate. "Up to 80% efficiency gains" is not a claim at all.
- **"How does it handle a document that is Arabic and English on the same page?"** Not "do you support Arabic". The mixed page is the real case and it is where most systems degrade quietly.
- **"What would make you tell me not to do this?"** A firm that can describe the conditions under which it would advise against the project has thought about the conditions under which it works.
- **"Who owns what at the end?"** The deliverables, the models, the prompts, the integrations, the data. Ambiguity here is discovered in year two, at renewal, which is the worst possible time to discover it.

8. The order that works

Five steps, and the order is the whole argument. Every failure mode in section 4 is a version of doing these in a different sequence.

Audit. Real data, real systems, the workflow as it runs rather than as documented. Two to four weeks is the market convention for a diagnostic and it is enough, provided the access is real.

Prioritise. A ranked map of opportunities against impact, effort and data readiness, including the explicit finding that a given case is not worth doing. Data readiness is the axis that gets left out and it is the one that decides which phase slips.

Redesign the process first. This is the step that gets skipped and it is the reason for most of the failure rate in section 3b. A broken process automated is a broken process that now runs faster and is harder to correct.

Implement in phases, against a baseline. The baseline is established in the audit, which is the only reason the phases can be judged. Flag rather than auto-act until the people doing the work have validated the flags.

Measure, against that baseline, and publish the result internally. Most AI projects in this market have no measured outcome, not because the outcome was bad but because nobody wrote down what it was before.

Audit before automate is the whole of it, and it is not a slogan about caution. It is a claim about where the cost is: building has become cheap, so the expensive decision is no longer how to build. It is what is worth building, in what order, and whether the organisation can absorb it.

9. What this document is not claiming

Four things, stated because a reader who has got this far is entitled to know where the edges are.

- **Not that AI will remove your back-office headcount.** The delivered results in section 3a are coverage and latency figures. What an operation does with recovered capacity is its own decision

and it is usually redeployment rather than reduction.

- **Not that the plant floor is the same problem.** Vision-based quality control, predictive maintenance and process optimisation are real and are not what this document is about. The back office is a narrower claim with better evidence behind it.
- **Not that the figures in 3a will reproduce on your material.** They are what happened on specific documents in specific operations. Whether they reproduce is what a diagnostic is for.
- **Not that this is urgent.** Nothing in the industrial back office is on fire. The paperwork will keep getting done. The argument for doing something about it is arithmetic, not alarm, and if the arithmetic does not work the answer is to leave it alone.

10. Where to start

Count the five things in section 6. That is the whole recommendation, and it works whether or not anyone is hired to help with it.

If the counting produces a number worth acting on, the smallest thing available is a two to four week diagnostic that examines real data in real systems and ends in a ranked map rather than a proposal. It can conclude that no AI work is warranted, and sometimes does.

Four calculators on the website put the arithmetic from section 6 into a form that takes about two minutes each, ungated, with no form in front of them: document load, search blind spot, quality-review coverage and voice authentication. They estimate. A diagnostic counts.

One Industry AI · Industrial AI for the Gulf back office · First Settlement, New Cairo, Cairo, Egypt · oneindustry.ai

Every figure in section 3a is from the practice's own delivery records for the engagement named, measured against the process that preceded it. No client is named in this document.