



March 2026, Frederik Bisbjerg

Using GenAI to increase trust and transparency in motor claims

A PRAGMATIC REAL-WORLD APPROACH

AI can materially raise transparency and trust in insurance by turning “trust” into an engineered outcome: cleaner data, faster and more consistent decisions, auditable workflows, and real-time customer visibility

In practice, that means using AI to (1) detect and explain anomalies and fraud early, (2) standardize decisions and communications so outcomes are predictable and defensible, and (3) embed governance directly into digital processes so compliance is demonstrable, not aspirational

The strategic shift is from “better automation” to “provable fairness and traceability” across the value chain, especially in claims where trust is won or lost



Crossroads ahead

The motor insurance industry in the GCC is approaching an inflection point. For decades, the fundamental economics of claims, the relationships that govern repair decisions, the opaqueness that surrounds pricing and quality, the manual processes that determine outcomes, have been sustained by a simple reality: there was no viable alternative.

The tools to capture, structure, and act on claims data at scale did not exist. The cost of the current model, while significant, was invisible because nobody had the means to measure it.

That is no longer true.

Artificial intelligence, and specifically the combination of large language models, computer vision, and structured data architecture, now makes it possible to do what the industry has talked about for a decade but never operationalized: **build claims processes where every decision is evidenced, every cost is benchmarked, every communication is logged, and every stakeholder, from the customer to the regulator, can see exactly what happened and why.**

This is not a technology dissertation. It is a paper about trust, and about the business case for making trust measurable. The insurers that move first will not just reduce fraud and leakage; they will build a structural advantage in customer retention, regulatory standing, and operational resilience that compounds over time. Insurers who wait and hope to sit this one out, will continue absorbing costs they cannot see, defending decisions they cannot evidence, and losing customers they never understood.

This paper lays out the case in eight chapters: why trust is breaking down, what genuine transparency looks like, how to build an AI trust stack with today's imperfect data, how governance keeps humans in control, where the highest-impact use cases are, what to measure, how to get started, and what to watch out for along the way.



Contents

CROSSROADS AHEAD	2
TRUST IS BREAKING DOWN	4
WHAT "TRANSPARENCY" MEANS IN INSURANCE	6
THE AI TRUST STACK: GETTING STARTED WITH WHAT YOU HAVE	9
GOVERNANCE AND COMPLIANCE-BY-DESIGN	13
HIGH-IMPACT USE CASES: THE TRUST STACK IN ACTION	15
IMPLEMENTATION ROADMAP	18
ABOUT THE AUTHOR	21
KPIS FOR TRUST AND TRANSPARENCY	23
COMMON FAILURE MODES AND GUARDRAILS	26

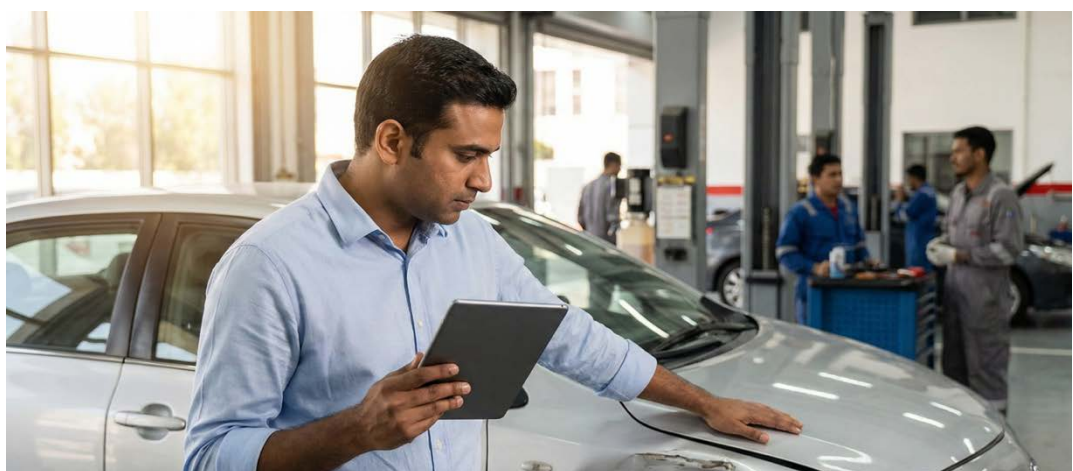


Trust is breaking down

Opaque decisions, inconsistent handling, slow outcomes, and rising fraud

Trust in GCC motor claims is breaking down because the economics and expectations of the market have modernized, while the repair ecosystem and many claims practices have not.

For decades, repair decisions have been mediated through an “old-guard” network: seasoned claims handlers relying on long-standing personal relationships with preferred garages and informal price norms.



That model can function when volumes are low, scrutiny is limited, and decision-making is largely trusted by default. It fails in today’s environment, where claim severity is rising, vehicle complexity is increasing, fraud is more organized, and regulators and customers expect provable fairness.

Even today, repair scopes and costs still originate as handwritten estimates, negotiated behind closed doors, and benchmarked against the same tight circle of workshops. This is not only inefficiency; it’s an evidence gap.

Insurers struggle to demonstrate that prices are market-consistent, that repair quality is verified, that decisions are free from bias or conflicts, and that leakage is actively controlled rather than retrospectively explained.



Key drivers eroding trust in motor claims, GCC context	
Evidence gap in pricing and quality Paper-based estimates, limited photo/parts traceability, weak benchmark data, inconsistent standards of repair validation.	Customer expectations shifting fast Customers now expect real-time visibility, predictable timelines, and clear explanations, not phone calls and ambiguity.
Network effects that entrench opacity “Comparable quotes” sourced from the same ecosystem, informal referral patterns, and limited competitive tension.	Regulatory and conduct pressure Higher expectations for documented decision rationale, complaint handling, service standards, and demonstrable control frameworks.
Rising severity and complexity Advanced driver-assistance systems, sensors, calibration needs, and higher parts costs increase variance and dispute risk.	Talent and scalability constraints Scarcity of modern claims analytics capability; heavy reliance on a few experienced individuals creates key-person risk and inconsistent outcomes.
Fraud and organized leakage Inflation of labor hours, parts substitution, add-ons, duplicate invoicing, staged damage, and referral economics that are hard to detect without data.	Fragmented supply chain with limited traceability Multiple intermediaries (recovery services, car rentals, parts suppliers, sub-contracted specialists) each add cost and complexity with little end-to-end visibility or accountability.

This is a clear and present balance-sheet, conduct, and brand risk problem compounding over time.

When decisions can't be evidenced end-to-end, leakage becomes structurally embedded in the loss ratio, disputes and reopenings rise, and cycle times stretch. That, in turn, drives higher complaint volumes, inconsistent outcomes across customers, and greater exposure in audits and regulatory reviews; the insurer cannot reliably demonstrate fairness, consistency, or effective controls.

Most importantly, the old model doesn't scale: it depends on individual judgment and relationships rather than repeatable processes and verifiable data.

In a market where customers expect transparency and regulators expect proof, insurers either industrialize trust through data, controls, and auditability, or they accept a persistent “trust discount” in profitability, compliance posture, and reputation.



Inconsistent, opaque claims operations has clear Board-level implications

- **Limited profitability:** severity inflation + hidden leakage becomes “normal,” reducing the ceiling on combined ratio improvement
- **Conduct risk:** inconsistent outcomes and weak decision traceability increase audit findings and complaint escalation
- **Operational resilience:** key-person dependency and informal practices create fragility and inconsistent quality
- **Customer experience:** low visibility and slow resolution erode retention and raise acquisition costs
- **Strategic positioning:** insurers fall behind peers who can prove fairness, speed, and control with evidence

What "transparency" means in insurance

Traceability, explainability, consistency, and customer visibility

Before exploring how AI can help, it's worth defining what genuine transparency looks like in motor claims. Not as it exists today, but as the standard the industry must be working towards. The reality, as outlined earlier, is that the insurance industry is far from this. But without a clear picture of the destination, incremental improvements risk being directionless.

In motor claims, transparency has three concrete dimensions; insights, explainability, and consistency.





Giving the customer insight into the repair

The biggest source of customer frustration is the lack of information. Once a vehicle enters the workshop, the policyholder typically enters an information vacuum: no clarity on the repair, what parts are being fitted, what quality checks are performed, or when the car will be returned.

Every repair already produces data: parts orders, labor records, inspection notes. The problem is that none of it reaches the customer. It sits in disconnected systems, in different formats, often in different languages.

Generative AI can ingest this data from workshops, parts suppliers, and quality checkpoints, and translate it into plain-language updates pushed directly to the policyholder: which garage has their car, what parts are being used, what stage the repair has reached, and what controls have been completed.

And equally important, the same information can be sent back to the insurer, creating an auditable record. What serves the customer also serves the claims file and the compliance function.

Customer-facing transparency:

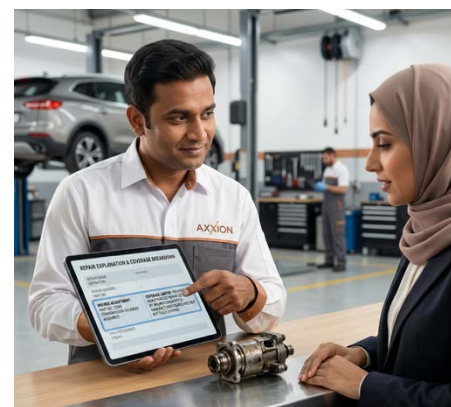
- Plain-language repair status updates pushed to the policyholder at each stage
- Visibility on the workshop handling the repair, including performance history
- Itemized parts information: OEM vs. aftermarket, supplier, and fitment confirmation
- Quality control evidence: photos, checklists, and sign-off records shared with the customer
- Every communication logged as part of the claims audit trail

Explaining decisions, not just communicating them

Transparency also means explaining outcomes, particularly when the insurer limits what it will cover.

A common dispute arises when pre-existing damage is identified during repair. Traditionally, a customer is told that certain damage "is not related to this claim" with little supporting evidence. From the customer's perspective, this feels arbitrary and frustrating.

AI-powered inspection tools offer a path forward. When a vehicle is assessed at the point of claim using structured photo capture and computer vision, its condition can be documented comprehensively. If the insurer can later show the customer a clear, time-stamped visual record that specific damage existed before the incident, the conversation shifts from confrontation to explanation.





Today, most claims lack this documented pre-repair baseline, and even where photos exist, they are often unstructured and inconsistent. But the principle is clear: explainability protects both sides. The customer gets a documented reason, not a vague rejection. The insurer gets a defensible position that holds up in disputes and regulatory reviews.

Explainability:

- AI-assisted condition assessments at FNOL, creating a verifiable baseline
- Visual evidence (time-stamped, geo-tagged) distinguishing claim damage from pre-existing wear
- Plain-language explanations linking each coverage decision to specific evidence
- A single evidence base serving the customer, the claims file, and any future dispute resolution

Consistency: from individual transparency to systemic trust

Customer visibility and explainability address individual claims. Transparency becomes truly powerful when it is systematic: when every claim follows the same verifiable path.

When every FNOL follows a standardized intake, every inspection uses the same AI-guided protocol, every repair is tracked against the same milestones, and every quality check is logged against the same criteria, the process becomes auditable by design. Regulators can be shown the exact process followed for any claim. Workshops know the standards they are held to. Insurance employees operate in a framework where every action is tracked, which has a natural deterrent effect on fraudulent or negligent behavior.

This is the "**sunlight effect**": when every participant knows that every action is recorded and auditable, behavior self-corrects. Fraud prevention becomes embedded in process architecture, not bolted on as retrospective detection.

This is the hardest dimension to achieve, because it requires not just technology but disciplined process adoption across the entire claims ecosystem. It demands that insurers, workshops, and service providers all commit to the same standards of data capture and transparency.

Today, that ecosystem alignment largely does not exist. Building it is a multi-year effort, but every step towards consistency compounds the value of the other two pillars.



Systemic consistency:

- Standardized processes across every claim, removing variance caused by individual judgment or relationships
- A complete audit trail demonstrable to regulators, reinsurers, and board governance at any time
- Workshop accountability through continuous, data-driven performance monitoring
- Natural fraud deterrence through tracking and tracing of every decision and transaction
- Operational resilience: the process no longer depends on key individuals or informal knowledge

These three pillars describe a target state, not today's reality. Most GCC motor markets operate far from this standard, and getting there will require sustained investment in data infrastructure, process discipline, and ecosystem collaboration. But the direction is clear, and the cost of *not* moving, rising leakage, regulatory exposure, customer attrition, and fraud, compounds every year.

The AI trust stack: getting started with what you have

Data integrity, decision guardrails, auditability, and customer communications

The transparency framework just described is ambitious. It is also, deliberately, a destination rather than a prerequisite. The most important message of this chapter is that the journey starts with small, practical steps, and even those first steps deliver measurable value.

Conversations about AI and data in insurance stall because the gap between the current state and the ideal feels insurmountable. Data is fragmented, inconsistent, mostly unstructured. Workshop systems are basic or non-existent. Internal processes vary by team, by handler, by day of the week.

And so the conclusion is: "we're not ready."

That conclusion is wrong. The AI trust stack is not an all-or-nothing investment. It is a set of layers that can be built incrementally, each one adding value on its own while creating the foundation for the next.



The four layers

At its simplest, the trust stack has four layers, each building on the one below.

4. Customer communications: Generative AI translating repair data into clear, proactive policyholder updates

3. Auditability: Automatic, end-to-end logging of every data point, flag, and decision as a byproduct of digital process

2. Decision intelligence: AI-powered triage, benchmarking, and anomaly detection, effective even on imperfect data

1. Data integrity: Minimum data model with quality gates, not perfect data, but consistent, structured, and verifiable

Layer 1: Data integrity. This is the foundation: making sure the right data is captured, in a consistent format, at the right points in the claims journey. It does not require perfect data. It requires a minimum data model, a defined set of data points that must be collected for every claim, and a clear specification of the format those data points must end up in.

The reality of today's claims ecosystem is that data arrives incomplete, inconsistent, and in every format imaginable: handwritten notes, unstructured emails, voice notes, photos with no labelling, PDFs that were never designed to be machine-readable.

Large language models can extract meaningful, structured information from these messy, imperfect sources. A handwritten estimate can be read and parsed. An unstructured email from a workshop can be interpreted and the relevant data points identified. Photos can be analyzed for damage type and severity. The LLM acts as the translation layer between the chaotic real world and the structured data model you need.

Once the AI has interpreted and extracted the data, traditional programming takes over, checks it for completeness and consistency, and routes it into the correct fields in the claims system. This ensures that the data entering your system is formatted correctly, that mandatory fields are populated, that values fall within expected ranges, and that nothing is lost or misclassified.

Generative AI for extraction from imperfect sources, deterministic programming for validation and structuring, is what makes a minimum data model achievable even in today's environment.



Layer 2: Decision intelligence. Once you have consistent data flowing in, AI start doing what it does best: pattern recognition, anomaly detection, and triage. Are the labor hours on this estimate in line with comparable repairs? Does the parts list match the documented damage? Has this workshop shown a pattern of inflated scopes? None of this requires a perfect dataset. AI is remarkably effective at surfacing inconsistencies and outliers even in imperfect data, and every flagged anomaly is an opportunity to learn, investigate, and improve.

Layer 3: Auditability. When data capture and decision support are digitized, auditability comes almost for free. Every data point collected, every AI flag raised, every human decision made in response is logged and traceable. This is the audit trail that regulators expect, that reinsurers value, and that protects the insurer in disputes. It is not a separate system to build; it is a natural byproduct of doing Layers 1 and 2 properly.

Layer 4: Customer communications. With structured data and an auditable process, generating clear, accurate customer updates becomes straightforward. Generative AI can translate the repair data into plain-language messages. The insurer can explain what is happening, why, and what comes next, because the underlying evidence exists. This is the layer customers see, but it only works because the layers beneath it are in place.

Starting small, but starting now

The minimum viable version of this stack is not a multi-year transformation program. It starts with defining the data you need for each claim, and insisting on getting it.

This means a structured FNOL intake with mandatory photo capture. A standardized estimate format (or at least capture of required data, Ai can standardize it). Digital confirmation of parts used and work completed. Basic milestone tracking from assignment to delivery. These are not unreasonable requirements. They are the kind of data that any competent workshop already has; it simply hasn't been asked for in a structured, verifiable way.

The enrichment comes over time. As data accumulates, benchmarks emerge. As benchmarks emerge, anomalies surface. As anomalies are investigated, processes tighten. Each cycle builds on the last. The insurer that starts capturing structured data today will, within months, have a dataset that enables meaningful AI-driven insights, even if the starting point was a blank page.

API connectivity accelerates this, connecting the insurer's claims platform to workshop management systems, parts databases, and inspection tools so that data flows automatically rather than being re-keyed or uploaded manually. But APIs are an accelerator, not a prerequisite. The first step is simply agreeing on what data is needed and collecting it consistently.



The workshop conversation

This is where insurers, and the partners working alongside them, must be willing to hold the line.

The workshops that dominate the current ecosystem are accustomed to operating on their own terms: handwritten estimates, informal communications, and limited accountability for the data they provide. Introducing structured data requirements will feel like a change, and change creates friction.



Requiring a workshop to submit structured photos, use a standardized estimate format, and confirm parts digitally is not a technology burden. Most of it can be done with a smartphone and a simple portal. It is a process change, not a capital investment.

The conversation with workshops needs to be clear and direct: "If you want to work with us, this is the data we need, and this is the standard of evidence we require." This is not adversarial; it is professional. It protects the workshop too, by giving them a documented record of the work they have done and the quality they have delivered.

The trust ecosystem starts with a simple requirement: verifiable data, collected consistently, at every stage of the claim.

What the minimum data model looks like:

- **At FNOL:** Structured photo set (minimum angles defined), standardized damage description, vehicle identification and condition baseline
- **At estimate:** Digital estimate in a comparable format, itemized parts and labor, benchmark-ready pricing
- **During repair:** Milestone updates (parts ordered, repair started, quality check, completion), parts confirmation (OEM/aftermarket, supplier)
- **At delivery:** Completion photos, quality inspection sign-off, customer confirmation
- **Throughout:** Every data point time-stamped, geo-tagged where relevant, and logged to the claims audit trail



Every journey starts with a first step, and in motor claims, that step is smaller than most insurers think. It is not about building a perfect AI platform on day one. It is about deciding what data matters, insisting on getting it, and using simple tools to start capturing it consistently. The AI trust stack grows from that foundation, layer by layer, and the value compounds with every claim processed.

The insurers who recognize that they cannot solve the transparency, fraud, and trust challenges described in this paper on their own will look for partners who can. The right partner brings the technology, the process discipline, and the workshop management capability to operationalize the trust stack from day one, so the insurer can focus on what it does best: underwriting risk and serving customers.

Governance and compliance-by-design

Humans stay in control. The system proves it.

A reasonable concern with any AI-driven process is: who is actually making the decisions? The answer, in a properly designed trust stack, is straightforward: humans do. AI surfaces information, flags anomalies, and accelerates workflows. But at every critical juncture, deterministic logic governs what happens next, and that logic is designed to keep humans in the loop where it matters.



In practice, this means threshold-based approval gates built into the claims workflow. If an estimate exceeds a defined value, the process pauses and routes to a senior handler for review. If an AI flag identifies a potential fraud indicator, the claim is escalated to an investigator, not auto-declined. If parts costs deviate from benchmark ranges, a human reviews before authorization. These are not AI decisions; they are rules-based checkpoints, coded in traditional programming, that determine when the process continues automatically and when it stops for human judgment.



This is the same principle described in the data layer earlier: AI handles extraction and pattern recognition; deterministic programming handles control logic. The AI will not approve claims, authorize a payment, or reject a repair scope. It informs the human who does.

Compliance by design

The deeper benefit of this architecture is what it gives the insurer from a regulatory and governance perspective. When every step in the claims process is digitized and logged, every data point time-stamped, every AI flag recorded, every human decision captured with its rationale, the insurer is compliant by design rather than by retrospective effort.

Consider what this means in practice. The regulator asks to see the decision trail for a specific claim: it is available in seconds, end-to-end, from FNOL to settlement. An internal audit reviews pricing consistency across a portfolio of repairs: the data is structured, comparable, and exportable. A customer disputes a coverage decision: the evidence base, including the vehicle condition at FNOL, the AI assessment, and the handler's documented reasoning, is already assembled.

This is a fundamentally different posture from the one most insurers occupy today, where compliance depends on the quality of individual file notes, the memory of the handler involved, and the ability to reconstruct a decision trail after the fact. In a digital, AI-supported process, the audit trail is not something you build when someone asks for it. It exists because the process itself creates it.

What compliance-by-design delivers:

- Threshold-based approval gates that pause the process for human review at defined trigger points
- Every decision, data point, and AI flag logged with timestamps and user attribution
- Full claim traceability recoverable in seconds, not days, for any regulator, auditor, or dispute
- Segregation of duties enforced by system logic, not by policy documents alone
- A defensible record that the insurer followed its own processes, consistently, for every claim

For financial services institutions operating under increasing regulatory scrutiny, this is not a nice-to-have. It is the difference between being able to demonstrate control and hoping that control existed. The trust stack builds the governance infrastructure that regulators, Boards, and reinsurers increasingly demand.



High-impact use cases: the trust stack in action

Claims, underwriting, and customer service

This chapter shows what it looks like when you put that architecture to work across the three areas where trust is most visibly won or lost.

Claims: evidence-first orchestration

This is the primary use case and the one that most directly addresses the trust breakdown described earlier. The objective is to replace relationship-based handling with an evidence pipeline that produces three things the old model cannot: a consistent baseline of vehicle condition so disputes don't become opinion battles, comparable repair scope and pricing across workshops so benchmarking is real rather than social, and a machine-readable claim record where every change to scope, parts, labor, and approvals is traceable.

In practice, this means the trust stack runs end-to-end across the claim. At FNOL, computer vision classifies and tags damage images while LLMs extract structured data from whatever format the workshop submits. Deterministic logic validates and routes. During assessment, AI benchmarks the estimate against comparable repairs and flags anomalies: labor hours above peer norms, parts lists that don't match documented damage, pricing inconsistencies, patterns associated with a specific workshop. Not decisions; but flags for review, routed to the right person at the right time through the approval gates.

Throughout the repair, milestones are tracked, parts confirmed, and quality checks logged. At every stage, the same structured record powers both the customer-facing updates and the insurer's audit trail. The strategic point is worth repeating: you are not automating claims. You are industrializing evidence, so that leakage, fraud, and inconsistency become visible and actionable.

What evidence-first claims orchestration delivers:

- Structured extraction from any input format (photos, PDFs, handwritten estimates, emails) into a comparable data model
- Real-time anomaly detection: scope inflation, parts substitution, labor hour outliers, repeat patterns by workshop
- Fraud signals assembled as an evidence pack with explainable pointers, not accusations
- Proactive customer updates derived from the same auditable record the insurer relies on
- Every change to scope, parts, approvals, and settlement logged and traceable



Underwriting: explainable risk and pricing governance

The underwriting version of opaqueness is inconsistent decisions, undocumented judgment, and weak traceability of why a quote moved. The same trust stack principles apply, adapted to a different workflow.

AI acts as the document intelligence engine. Underwriting submissions arrive as PDFs, spreadsheets, emails, and broker notes. LLMs extract insured details, vehicle specifications, usage, claims history, coverage requested, and special terms into a structured submission record. Deterministic logic then enforces a minimum submission standard: no quote progresses without required fields, and missing items generate a structured request list back to the broker.

At the decision intelligence layer, AI surfaces risk signals and governance triggers. It spots outliers in frequency and severity patterns, highlights inconsistencies between renewal and new submission data, and flags incomplete disclosures. For pricing, it suggests placement within corridors based on comparable risk profiles and forces structured rationale capture when a human deviates from norms. Importantly, traditional actuarial pricing remains intact. AI's job is to ensure the inputs are complete, the deviations are visible and justified, and the entire quote-to-bind journey is reconstructable.

The governance principles apply directly: referral rules encoded deterministically, mandatory justification for overrides, and no autonomous binding. AI cannot bind or alter coverage; it can only propose and explain.

What explainable underwriting delivers:

- Structured extraction from submissions, reducing manual re-keying and catching conflicts across documents
- Risk signals and anomaly detection surfaced before the underwriter commits
- Pricing governance: deviations from corridors are visible, justified, and logged
- Broker and customer-ready explanations of what drives terms and what actions could improve them
- Full quote-to-bind traceability for conduct reviews and audits

Customer service: proactive transparency

We described the escalation cycle that erodes trust: slow updates lead to frustration, frustration leads to complaints, complaints lead to disputes and reopened claims. This use case breaks that cycle by making transparency an always-on service rather than a reactive exercise.



The key design principle is that the customer-facing assistant reads from the same structured claim record created in the claims use case. It does not invent; it retrieves and summarizes. When a customer asks "where is my car," the assistant can answer with facts: the workshop handling the repair, the current stage, the parts ordered, the estimated completion. When a coverage decision needs explaining, it references the evidence pack rather than improvising.

AI monitors for complaint risk: sentiment patterns across interactions, SLA breaches, stuck milestones, high-friction moments like scope reductions or parts delays. When risk is detected, the system proposes proactive actions: an expectation-resetting update to the customer, an escalation to a supervisor with a summary and evidence pack, or a follow-up trigger to the workshop.

Governance here is especially important because customer-facing AI fails visibly when it hallucinates or over-commits. Hard boundaries ensure the assistant cannot approve claims, promise settlement dates, or interpret coverage beyond policy rules. Low-confidence answers route to a human. Every output is logged and becomes part of the claims audit trail.

What proactive customer service delivers:

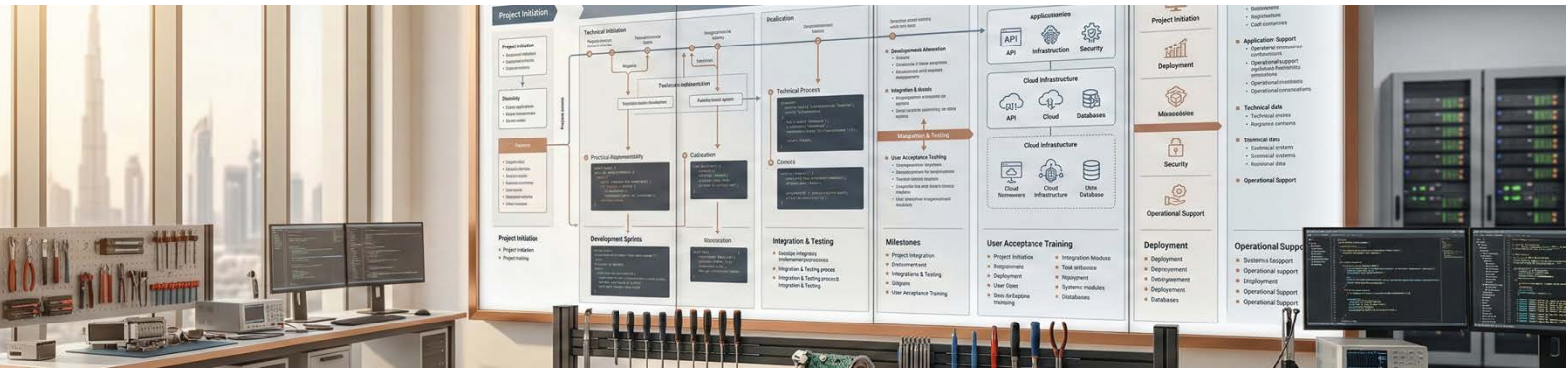
- Accurate, evidence-grounded answers to customer queries drawn from the structured claim record
- Proactive updates pushed at key milestones without the customer needing to chase
- Complaint risk detection and early intervention before frustration escalates
- Consistent quality of explanation regardless of channel, time, or handler availability
- Every interaction logged as part of the compliance and audit framework

These three use cases are not independent projects. They share the same data layer, the same governance framework, and the same audit architecture. An insurer that builds the trust stack for claims automatically creates the foundation for explainable underwriting and proactive customer service. The investment compounds across functions, and the evidence base grows with every transaction processed.



Implementation roadmap

Start small. Start now. Scale on evidence.



Insurers reading this paper will recognize the problems described. They have known about them for years. The reason nothing changes is not lack of awareness; it is inertia. The current model works in the narrow sense that claims get paid, workshops get used, and customers mostly don't leave. The pain is diffuse - leakage spread across thousands of claims, fraud that is never detected, quality failures that surface as complaints months later - rather than acute. There is no single crisis that forces action, so the default is to keep going.

The cost of the current model is invisible precisely because nobody is measuring it. The roadmap below is designed to make the first step so small that it is harder to say no than yes.

Phase 1: See what you've been missing (days 1-90)

This is not a transformation program. It is a proof of concept on a contained subset of claims, a single line of business, a specific region, or a defined volume.

Implement structured FNOL data capture with mandatory photo sets. Apply basic AI extraction to convert whatever the workshops submit into structured data. Run the minimum data model described earlier. The only question you are answering is: *what does the data reveal that you couldn't see before?*

The answer will be significant. Patterns in estimate inflation, inconsistencies between damage photos and repair scopes, pricing variance across workshops, missing milestones - all of it becomes visible for the first time, simply because you started collecting the data in a structured way.

Low cost. Low risk. High signal.



Phase 2: Benchmark and measure (months 4–9)

With structured data flowing, now turn on the decision intelligence layer. AI begins benchmarking estimates, flagging anomalies, and triaging claims by complexity and risk. Start tracking KPIs - not all of them, but the ones that matter most: FNOL evidence completeness, flag precision, estimate scope accuracy, cycle time by stage.

This is also the phase to begin the workshop conversation in earnest. Set the data standards. Communicate them clearly. Monitor compliance. The workshops that engage will see the benefit of a documented, defensible record of their work. Those that resist will make themselves visible.

The first financial signals appear here: leakage identified and quantified, scope anomalies flagged before approval, cycle time patterns that explain cost overruns. These are the numbers that justify Phase 3.

Phase 3: Operationalize (months 9–18)

Scale the trust stack across the participating portfolio. Customer communications layer goes live: proactive updates, evidence-based explanations, milestone tracking visible to the policyholder. Governance framework fully embedded: approval gates, override logging, compliance-by-design audit trail.

Scale decisions beyond this point are driven by evidence from Phases 1 and 2, not by projections or assumptions.

Roadmap summary:

- **Phase 1 (90 days):** Structured FNOL capture + AI extraction on a contained claim subset. Answer: what can we now see that we couldn't before?
- **Phase 2 (months 4–9):** Decision intelligence active. KPI tracking live. Workshop data standards enforced. First financial impact quantified.
- **Phase 3 (months 9–18):** Full trust stack operational. Customer comms live. Governance embedded. Scale based on evidence.



The harsh reality

Most GCC insurers do not have the internal technology team, workshop management capability, or process design expertise to build this themselves. And they should not need to. The realistic path for most is a TPA partner that brings the trust stack ready-made: the insurer provides the portfolio, the partner provides the infrastructure, the data discipline, and the workshop accountability.

You're not outsourcing claims. You're operationalizing the trust architecture described in this paper with a partner whose entire operating model is built around evidence-based claims management, structured data, and transparent processes. The insurer retains oversight, governance, and strategic control. The partner delivers the execution capability.

The cost of waiting

Every month without structured data is another month of invisible leakage, undetected fraud, and unreported quality failures. That cost does not sit in a single line item; it is embedded in every loss ratio that cannot be fully explained, every complaint that could have been prevented, and every regulatory review that depends on reconstructing decisions after the fact.



About the Author

[Frederik Bisbjerg](#) is Co-founder and Managing Director of Axxion Claims Settlement Services LLC, the UAE's first dedicated motor claims third-party administrator, where he is building a compliance-by-design claims operating system with AI governance at its core.

His career spans more than two decades of insurance leadership across the MENA region, including roles as CEO of Al Wathba Insurance, Chief Transformation Officer at AXA Global Healthcare, Senior Vice President of Digital Transformation and Innovation at Daman National Health Insurance Company, and Executive Vice President at Qatar Insurance Group.

Before moving into executive roles, he spent several years at a top-tier management consulting firm, where he developed the habit of building alliances between business partners that had not previously thought to work together.

He also serves as Head of MENA and Digital Transformation specialist at The Digital Insurer, where he is a founding member of the world's first mini-MBA in Digital Insurance and lectures on strategy, transformation, big data, and technology architectures.

Bisbjerg is the author of the best-selling *Insurance_Next*, a practical guide to transforming incumbent insurers into flexible, resilient organizations ready for the post-COVID, generative-AI era. The ideas in this white paper grew directly from his experience: the gap between what AI can technically do in claims and what governance structures actually exist to make it safe, auditable, and worth an insurer's trust.

About Axxion Claims Settlement Services

[Axxion Claims Settlement Services](#) is a Dubai-based end-to-end motor claims management company and the UAE's first dedicated motor TPA. Axxion manages the full claims lifecycle for insurance partners, from first notification of loss through repair coordination, quality control, and settlement, operating on a six-layer claims architecture designed around regulatory compliance, data integrity, and AI-augmented decision-making.

Axxion is part of the Skelmore Group, a diversified automotive and insurance services group founded in Toronto in 1994. The group operates across North America and the Middle East with approximately \$650 million in revenue and 4,000 employees, spanning multi-brand automotive aftermarket services, retail and wholesale distribution, and luxury automotive.





APPENDIX



KPIs for trust and transparency

What to measure, and why it matters

Transparency is only meaningful if it can be measured. The following KPIs give insurers a practical scorecard for tracking whether the trust stack is delivering results across the full claims lifecycle. None require perfect data to start tracking; most can be baselined from existing operations and improved as the data foundation matures.

Data integrity

FNOL evidence completeness rate

% of claims with the minimum required photo set and mandatory fields captured at first contact. Predicts dispute rate, rework, and cycle time downstream.

First-time-right data rate

% of claims requiring no follow-up for missing or incorrect core data (VIN, driver, incident details, photos). Direct driver of cycle time and cost-to-serve.

Workshop data compliance score

Quality-weighted compliance with the required data model: structured photos, estimate format, milestone updates, parts confirmation. The measure of ecosystem adoption.

Decision intelligence

Flag precision and yield

% of AI flags that lead to confirmed leakage, fraud, or adjustment. Proves the system is finding real issues, not generating noise. Yield measured as value recovered or avoided per 1,000 claims.

Estimate scope accuracy

Gap between initial estimate and final settled scope (excluding genuine hidden damage supplements). Indicates quality of damage assessment and workshop discipline.

Supplement rate and severity

Frequency and size of supplements after repair starts. High rates signal weak early evidence capture or strategic under-scoping by workshops.

Leakage avoided / recovered

Monetary value of scope reductions, pricing corrections, duplicate detection, and recoveries triggered by AI-surfaced anomalies. The core financial ROI measure.



Auditability and governance

Decision traceability score

% of key decisions with a complete rationale chain: evidence → rule or policy basis → approver. The strongest governance KPI in this framework.

Override rate (and justified override rate)

How often handlers override AI flags or benchmark guidance, and whether they document why. Signals drift, bias, or operational mismatch.

Approval gate compliance

% of claims exceeding thresholds that properly trigger review and approval. Proves deterministic control logic is being followed.

Audit retrieval time

How quickly a full decision trail can be produced for any given claim. In a compliance-by-design environment, this should be seconds, not days.

Regulatory and audit findings

Number and severity of findings in internal audits, regulatory reviews, and reinsurer assessments. Fewer findings confirm that controls are working.

Customer transparency and experience

Proactive update coverage

% of claims where the customer received milestone updates without asking. The direct measure of whether the "information vacuum" is being solved.

Inbound status-chasing contact rate

Customer contacts per 1,000 claims that are pure "where is my car?" inquiries. If transparency is working, this drops sharply.

Time-to-first meaningful update

Hours from FNOL to first customer communication containing concrete next steps. Sets the trust tone early.

Dispute rate and resolution time

How often customers dispute scope or coverage decisions, and how fast those disputes are resolved. More precise than complaint rate as an explainability measure.



Complaint rate (per 1,000 claims)

The broadest measure of customer trust. A declining rate indicates that transparency, explainability, and proactive communication are working together.

Customer NPS / satisfaction score

A lagging but important indicator. Improvements should translate into measurable shifts in sentiment over time.

Repair quality and operational performance

End-to-end repair cycle time (FNOL to vehicle delivery)

The single most visible operational measure for both insurers and customers. Decomposed by stage (assignment, estimate, parts, repair, QC, delivery) to identify where AI and process discipline are moving the needle.

Comeback rate (repair rework)

% of repairs returning for defects within 30/60/90 days. Connects transparency directly to actual repair quality and reduces future claims cost.

Quality checkpoint pass rate

% of repairs passing quality control on first inspection. Operational proof of workshop standards and discipline.

Parts lead-time variance

Variation between expected and actual parts arrival. A major driver of customer frustration and rental cost overruns.

Rental / loss-of-use days per claim

Rental days and cost per claim. Hard financial impact directly tied to repair delays and process efficiency.

Financial and portfolio (Board-level)

Leakage ratio

Estimated leakage as % of paid losses (scope inflation, pricing variance, parts substitution). Tracks structural improvement, not just individual wins.

Cost-to-settle per claim



Operational cost (handling touches, vendor interactions, rework) per claim. Structured workflows and AI triage should reduce this progressively.

Claims severity variance

Variance around expected severity for comparable claim types. Tightening variance indicates consistency and control across the portfolio.

Common failure modes and guardrails

What can go wrong, and how to prevent it

AI in insurance fails predictably when governance is weak. Building the trust stack without acknowledging these risks would be naive. These are the failure modes we see most often, paired with the guardrails that prevent them.

Hallucinated explanations

Generative AI can produce outputs that read convincingly but are factually wrong, fabricating rationale, inventing data points, or citing evidence that doesn't exist. In a claims context, this is dangerous: a hallucinated explanation sent to a customer or used in a dispute becomes a liability, not an asset.

The mitigation is architectural, not behavioural. Customer-facing and decision-support AI must operate in "grounded generation" mode: every output must reference and be derived from structured claim data, and the system must be designed so that if the underlying data doesn't exist, the AI cannot fill the gap with invention. Free composition should never be permitted for explanations, coverage decisions, or audit-facing content.

Alert fatigue

When AI generates too many flags, most of which turn out to be noise, handlers learn to ignore them. This is worse than having no flags at all, because it creates the illusion of oversight while real anomalies pass through unchallenged.

The mitigation is continuous measurement and tuning. Flag precision (the percentage of flags that lead to confirmed findings) must be tracked as a KPI. Thresholds should be adjusted regularly based on outcomes, and signals with consistently low precision should be retired or redesigned rather than left running.

Automation bias

The opposite of alert fatigue: handlers trust AI output too readily and stop applying their own judgment. When AI says "no anomaly detected," the handler assumes the claim is clean. When



AI suggests a repair scope, the handler approves without scrutiny. Over time, the human-in-the-loop becomes a rubber stamp.

The mitigation is structural. Threshold-based approval gates force genuine review at defined trigger points, and override decisions must be logged with documented rationale. Periodic audit sampling should specifically test whether handlers are engaging critically with AI outputs or simply confirming them.

Privacy leakage

In a system where multiple stakeholders (insurer, workshop, customer, regulator) access the same underlying data, the risk of exposing information to the wrong party is real. A customer sees internal fraud notes. A workshop accesses another workshop's pricing. An AI-generated summary includes data from an unrelated claim.

The mitigation is role-based access control with deterministic permissioning: each stakeholder role sees only the data fields and outputs they are authorized to see, enforced by system logic rather than user discipline. Data segregation between stakeholders must be designed into the architecture from the start, not added as a layer afterwards.

Workshop data gaming

Workshops that are required to submit structured data will, predictably, find ways to meet the letter of the requirement without meeting its intent. Photos that technically exist but show nothing useful. Estimates that are formatted correctly but contain inflated or fabricated line items. Milestone updates submitted in bulk after the fact rather than in real time. Simple completeness checks will not catch this.

The mitigation is quality-weighted compliance scoring that assesses not just whether data was submitted, but whether it is accurate, timely, and internally consistent. Cross-referencing across data points (do the photos match the estimate? do the parts listed match the damage documented?) is where AI adds real value in detecting gaming behavior.

Model drift

AI models are trained on historical data. As fraud patterns evolve, repair costs shift, vehicle technology changes, and workshop behavior adapts, the models become less accurate. A triage model trained on last year's data may misroute claims this year. A benchmark that was valid six months ago may now be outdated.

The mitigation is treating models as living systems, not finished products. Model versioning ensures you know which version produced which output. Performance monitoring tracks whether flag precision, triage accuracy, and benchmark relevance are holding or degrading. Scheduled recalibration, using the new data the trust stack itself generates, keeps models current.



Change management failure

This is the most common failure mode and the least technical. The technology works, but the people and processes don't follow. Handlers revert to old habits because the new workflow feels slower or less familiar. Workshops disengage because the data requirements feel burdensome and nobody enforces them. Management loses patience because the benefits take longer than expected to materialize.

The mitigation is phased rollout with demonstrated quick wins before scaling. Handlers need training not just on how the system works, but on why it exists and what it changes for them. Workshops need to see that compliance is monitored and that the standards are non-negotiable. And leadership needs realistic timelines with interim metrics that show progress before the full trust stack is operational.

None of these failure modes are unique to insurance, and none are unsolvable. They are predictable, and the guardrails are well understood. The difference between a trust stack that delivers and one that disappoints is almost always governance discipline, not technology capability.