

**SHIFT**

# **ARISE: A standard framework for AI agent autonomy in insurance**

Authored by Shift Technology | May 2026

[www.shift-technology.com/en-gb](http://www.shift-technology.com/en-gb)

# Executive summary

Insurance is at an inflection point. AI agents are moving from pilots to production, from assistants to autonomous decision-makers — and the industry has no shared language to describe what that actually means. Without common definitions, insurers cannot set realistic deployment expectations, vendors cannot differentiate their capabilities honestly, and regulators cannot establish proportionate oversight. The result is a market saturated with competing claims and a profession struggling to distinguish a chatbot from a fully autonomous claims engine.

Shift Technology proposes to fill that gap with ARISE — a standard framework for AI agent autonomy in insurance. The name is both an acronym and an organising idea. The five levels — Answers, Recommends, Initiates, Solves, Exceeds — describe the journey every insurer will take as AI agents grow in capability and earn greater trust. Drawing on more than a decade of deploying AI in production environments at carriers representing over 350

million policyholders globally, and modeled on the SAE International J3016 standard that brought lasting clarity to the autonomous-vehicle industry, ARISE gives the insurance sector the precise, vendor-neutral vocabulary it has been missing: a way to evaluate, procure, and govern AI agent capabilities that is grounded in operational reality rather than marketing aspiration.

This white paper presents the full ARISE framework — with precise definitions and real-world examples for each level — maps Shift Technology’s current production deployments and product roadmap to each level across auto, property, workers’ compensation, and travel lines, and makes the case that a standard autonomy taxonomy is a prerequisite for the industry’s responsible adoption of agentic AI. We offer ARISE not as a proprietary tool but as a contribution to the infrastructure of accountable AI in insurance — a standard we invite carriers, vendors, regulators, and analysts to adopt, stress-test, and build upon.

## 1. Why insurance needs an autonomy standard

### The Precedent: SAE J3016 and the Automotive Industry

In 2014, SAE International published J3016, “Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles.” The standard defined six levels of driving automation — from Level 0 (No Automation) to Level 5 (Full Automation) — and became the de facto global reference for automakers, regulators, insurers, and consumers alike. Its success rested on three properties: precision (each level had a clear, testable definition), universality (it applied regardless of manufacturer or platform), and parsimony (six levels captured meaningful capability thresholds without needless granularity).

The aviation sector followed a similar path. The FAA and EASA have long used structured automation levels to govern cockpit systems, from manual flight through autopilot to highly automated envelope protection. In industrial robotics, the ISO 8373 standard defines levels of robot autonomy that anchor procurement specifications, safety assessments, and insurance underwriting for factories worldwide.

In each case, the industry arrived at a point where rapid technological advancement outpaced shared understanding — and a formal taxonomy restored clarity. Insurance AI is at precisely that inflection point today.

## The Current State of Insurance AI: A Vocabulary Problem

McKinsey & Company estimated in 2023 that AI-enabled insurance use cases could generate £823.7 billion in value annually across the global industry — yet adoption remains uneven and poorly measured. One reason is definitional ambiguity. A carrier’s RFP for an “AI claims agent” might attract responses from vendors offering anything from a simple chatbot to a fully autonomous straight-through processing engine. Without a shared taxonomy, meaningful comparison is impossible.

Industry surveys reinforce the problem. A 2024 Deloitte survey of insurance executives found that while 79% described AI as a strategic priority, fewer

than 30% reported having a defined framework for evaluating agent autonomy or establishing human oversight requirements. The gap between aspiration and governed deployment is wide — and a standard taxonomy directly addresses it.

**The consequences of definitional ambiguity are practical and financial.** Insurers overpay for AI features they do not use. Regulatory examinations of AI systems produce inconsistent findings because examiners lack consistent reference points. And the reputational risk of an AI agent acting at a higher autonomy level than the organisation intended — or believed — is significant. A standard taxonomy converts these risks into manageable, measurable decisions.

## 2. The Shift levels of autonomy framework

The Shift Levels of Autonomy (SLA) framework defines five discrete levels of AI agent capability in insurance. Together, the level names — Answers, Recommends, Initiates, Solves, Exceeds — form the acronym ARISE. Like SAE J3016, each level is defined by the degree of human involvement required and the complexity of judgment the agent exercises independently. The levels are cumulative: an agent operating at L3 implicitly possesses the capabilities of L1 and L2.

### Level 1 — Answers (A): Intelligent Information Retrieval

At L1, the agent responds to natural-language questions by synthesising information from structured and unstructured sources — policy documents, claims records, regulatory databases, and case history. The agent does not recommend actions; it informs. This level corresponds broadly to what SAE J3016 calls Level 1 in the driving context: a single automated function that assists the human operator without replacing any aspect of decision-making.

**Insurance example:** A claims handler asks, “Is storm damage to the roof covered under this homeowners policy given the policyholder’s deductible structure?” The agent retrieves the relevant policy sections, applies coverage logic, and returns a plain-language answer with citations.

**Production evidence:** Shift’s L1 agents are deployed in production across auto, property, workers’ compensation, and travel lines today. Deployments at this level deliver approximately 10% efficiency gains by eliminating manual policy lookups and reducing handler research time.

### Level 2 — Recommends (R): Situational Analysis and Recommendation

At L2, the agent moves beyond answering questions to proactively analysing a situation and recommending a sequence of next best actions — which is why this level is called Recommends. It synthesises multiple data streams — claim details, documents, photographs, third-party data — applies jurisdictional rules, and

surfaces a prioritised action plan. The human handler retains full decision authority and must act on each recommendation. This parallels SAE L2 partial automation, where the vehicle handles multiple functions simultaneously but the driver remains responsible for monitoring and override.

**Insurance example:** After FNOL on a multi-vehicle auto accident, the agent advises the handler to initiate the repair process, order a police report in parallel, and contact identified witnesses before the 72-hour

evidence-preservation window closes — all ranked by time sensitivity.

**Production evidence:** L2 Recommends capability is live in production across Shift’s auto liability, property building and content, and workers’ compensation coverage modules. Organisations operating at L2 realise approximately 20% efficiency gains and a 1% improvement in indemnity outcomes, as sharper investigator focus drives more accurate early-stage assessments.

**Table 1: The Shift levels of autonomy (SLA) framework for insurance AI agents**

Level	Name	Definition	Example: claims processing	Efficiency gain	Indemnity impact
L1	Answers	The AI agent responds to direct questions, retrieving and synthesising relevant information from policy documents, claims records, and regulatory sources.	“Is this claim covered under the policyholder’s comprehensive auto policy?”	10%	—
L2	Recommends	The agent analyses the full situation — claim details, documents, jurisdiction — and recommends best next steps with clear rationale.	“Initiate the repairs process and, in parallel, order a police report and contact identified witnesses.”	20%	1%
L3	Initiates	The agent initiates all required checks, pre-fills decision parameters, and presents a fully validated action package ready for one-click human approval.	“Here are all checks performed; the recommended payment amount and all correspondence have been pre-filled — awaiting your approval.”	30%	1%
L4	Solves	The agent acts end-to-end without human intervention, achieving 99%+ accuracy by applying contractual, regulatory, and insurer-specific logic consistently at scale.	The agent applies all contractual, regulatory, and insurer-specific logic and makes fewer than 1% errors.	50%	2%
L5	Exceeds	The agent not only operates autonomously but surpasses the outcomes of the top 1% of human performers, proactively identifying process inefficiencies and deviating intelligently to optimise results.	The agent identifies where an insurer’s standard process is sub-optimal and takes initiative to deviate, producing the best possible outcome for both insurer and policyholder.	80%	3%

### Level 3 — Initiates (I): Human-Validated Autonomous Execution

L3 represents the critical threshold between assistive and agentic AI — and the name Initiates captures this precisely: the agent initiates the full workflow and prepares every element for execution, but a human expert provides final authorisation. The agent completes all required checks — coverage verification, fraud screening, liability assessment, reserve calculation, document validation — and packages the results with pre-filled decisions and correspondence ready for a single human approval gesture. Human oversight is preserved but reduced to the final authorisation step. This mirrors SAE Level 3 conditional automation, where the system manages the full driving task in defined conditions but requires a human to be available to intervene.

**Insurance example:** The agent presents: “Coverage confirmed. Estimated repair cost \$4,840 — within policy limits. No fraud indicators. Vendor authorised. Pre-filled payment authorisation and customer communication are ready. One click to approve.”

**Roadmap status:** L3 Initiates capability is targeted for 2026 deployment across auto subrogation, auto and property injury management, and workers’ compensation injury management modules. At this level, organisations achieve approximately 30% efficiency gains and a further 1% improvement in indemnity outcomes — reflecting the combined effect of faster approvals and more consistent application of decision logic.

### Level 4 — Solves (S): Full Straight-Through Processing at 99%+ Accuracy

At L4, the agent executes the full workflow end-to-end — intake, investigation, decision, and disbursement — without requiring human review for each transaction. The system applies contractual terms, regulatory requirements, and insurer-specific business rules consistently at scale, achieving 99%

or higher decision accuracy. Human oversight shifts from transaction-level approval to portfolio-level audit and exception management. This parallels SAE Level 4 high automation: the vehicle manages all driving functions within a defined operational domain without human intervention.

**Insurance example:** For a glass repair claim that meets all straight-through criteria, the agent validates coverage, confirms the repair estimate against market benchmarks, authorises payment to the approved vendor, and sends the policyholder a confirmation —with no human touchpoint.

**Production evidence:** Shift has achieved L4 Solves capability in production for auto glass and APD repairs, property electronic devices, property building and content losses, workers’ compensation coverage, medical bill review, and travel trip interruption claims. L4 deployments deliver approximately 50% efficiency gains and a 2% indemnity improvement — the compound result of eliminating human touchpoints while maintaining 99%+ decision accuracy.

### Level 5 — Exceeds (E): Superhuman Performance and Process Innovation

L5 is the frontier of insurance AI. The agent not only operates autonomously but consistently outperforms the top 1% of human practitioners — not by following established processes more efficiently, but by identifying where those processes are sub-optimal and deviating intelligently to generate superior outcomes. L5 agents discover novel fraud patterns before they crystallise into losses, identify latent subrogation recoveries that experienced examiners miss, and dynamically re-route complex claims to accelerate resolution. This corresponds conceptually to SAE Level 5 full automation: capability exceeding the human envelope across all conditions.

**Insurance example:** The agent detects a pattern across 47 seemingly unrelated water damage claims filed over 18 months — all originating from properties managed by a single contractor — and proactively escalates the cluster as a potential organised fraud ring, generating a recovery opportunity that no individual examiner would have identified.

Roadmap status: Shift has L5 Exceeds capability targeted for 2026 in auto liability and property building

and content loss use cases, with further expansion across the product portfolio planned. L5 represents the peak of measurable impact: approximately 80% efficiency gains and a 3% improvement in indemnity outcomes, reflecting the agent’s ability to identify superior resolutions that even the most experienced human practitioners would typically miss.

### 3. Shift Technology: autonomy in practice

The SLA framework is not a theoretical construct — it is grounded in Shift Technology’s deployment experience across the world’s leading carriers. The table below maps Shift’s current product capabilities and planned roadmap items to each autonomy level, by insurance line and use case.

#### Key Milestones and Customer Evidence

- **L4 Solves addresses the financial materiality** of modern financial crime by uncovering hidden patterns and mitigating complex fraud at scale.
- **2x increase in subrogation recovery rates** and 33%+ faster time from FNOL to recovery in production subrogation deployments.
- **~70% claims coverage across all 50 U.S. states** via the Insurance Data Network (IDN), enabling cross-carrier pattern detection that underpins L5 Exceeds capabilities.

- **4x higher fraud investigation acceptance rates** versus industry baselines, a direct outcome of L2 Recommends and L3 Initiates precision improving investigator focus.
- **5x faster investigation times** for fraud cases supported by L3 Initiates pre-packaging.

These metrics are not aspirational. They are drawn from Shift’s current production deployments. The SLA framework was developed precisely because Shift has experienced firsthand the challenge of communicating these distinct capability levels to carrier executives, procurement teams, and regulators without a shared taxonomy to anchor the conversation.

**Table 2: Shift Technology product autonomy roadmap by insurance line and use case**

Line of business	Use case / scope	L1 Answers	L2 Recommends	L3 Initiates	L4 Solves	L5 Exceeds
<b>Auto</b>	Coverage & resolution – glass, APD repairs	In production	In production	In production	In production	—
	Coverage & resolution – theft, fire, total losses	2027	—	—	—	—
	Liability	In production	In production	In production	In production	2026
	Subrogation	In production	In production	2026	—	—
	Injuries (1st & 3rd party)	In production	In production	2026	—	—
<b>Property</b>	Coverage & resolution – electronic devices	In production	In production	In production	In production	—
	Coverage & resolution – building & content losses	In production	In production	In production	In production	2026
	Coverage & resolution – financial losses	2027	—	—	—	—
	Subrogation	In production	In production	2026	—	—
	Injuries (1st & 3rd party)	In production	In production	2026	—	—
<b>Workers comp</b>	Coverage	In production	In production	In production	In production	—
	Medical bills	In production	In production	In production	In production	—
	Loss of salary	2026	—	—	—	—
	Injury mgmt & return to work	In production	In production	2026	—	—
	Subrogation	In production	In production	2027	—	—
	Invalidity and death	2027	—	—	—	—
<b>Travel</b>	Coverage & resolution – trip interruption	In production	In production	In production	In production	—
	Coverage & resolution – rental vehicle & property	2026	—	—	—	—

## 4. Case study: ARISE in action

Shift's claims automation platform first achieved L4 Solves-level autonomy in 2020 for straightforward claim types. Since then, the scope has expanded steadily across increasingly complex lines of business. With the addition of agentic AI capabilities, Shift's claims agents now operate at L4 across virtually all claim types and lines of business. The two examples below illustrate the ARISE framework in action on the same bodily injury claim — first as it is deployed today in production at L3 Initiates, and then as a simulation of what L5 Exceeds capability would look like on that same claim with the same real data.

L3 Initiates is where Shift's bodily injury agents operate in production today. Human BI examiners remain engaged at structured decision points throughout the workflow, maintaining full oversight and authority over key judgments while the agents handle the analytical and operational work between those checkpoints.

### L3 Initiates in Production: Bodily Injury, Personal Auto Casualty

**Carrier and context.** A U.S. personal lines P&C carrier operating Shift agents for its personal auto casualty portfolio. On April 6, 2026, the carrier received a bodily injury (BI) demand package from a personal injury law firm seeking \$168,000 in compensation. The claim related to a crossroad collision that had occurred five months earlier in which the carrier's policyholder was assessed as at fault; the vehicle damage exposure had already been resolved. The demand package comprised 112 pages of claim statements, vehicle damage reports, medical records, and medical bills, and stipulated a 10-day response deadline.

#### Step 1 — Document intake and claim indexing.

The Shift agent ingested the demand package, classified each document type, extracted the demand conditions and deadlines, and linked the package to the underlying first-party claim record in the carrier's core system — presenting the indexed summary to the assigned BI examiner for confirmation before proceeding.

#### Step 2 — Liability analysis and evidence plan.

After retrieving the full claim history, the agent conducted a liability analysis and identified that the policyholder had reported the third party was exceeding the speed limit — a fact that, if substantiated, could reduce the carrier's liability share. The agent presented this finding and a recommended evidence-gathering plan to the BI examiner. Upon the examiner's approval, the agent collected the police report and traffic camera footage, building the evidentiary record that supported reducing the liability share from 100% to 75%.

#### Step 3 — Injury evaluation and contradiction

**detection.** In parallel, the agent reviewed all medical bills and records, benchmarking the claimed injuries against comparable settled claims and relevant case law. It also flagged an internal inconsistency: the claimant alleged inability to participate in a scheduled golf competition as evidence of injury severity, yet the medical records indicated no physical restrictions on that date. The agent surfaced both the damages range assessment and the contradiction to the BI examiner for review and sign-off before proceeding.

**Step 4 — Settlement package and one-click approval.** With the liability reassessment and damages analysis validated by the examiner, the agent assembled a fully documented settlement package — including the proposed counter-offer amount, supporting rationale, and pre-drafted correspondence — and presented it for one-click approval. The examiner approved. The agent submitted a settlement proposal of \$87,360 (52% of the original demand) on April 12; the third party accepted on April 14. Total elapsed time: eight days from receipt of the demand package, with the BI examiner engaged at four structured checkpoints rather than managing the claim manually throughout.

**Why this is L3 Initiates.** The Shift agent orchestrated every analytical and operational step of this complex claim — document ingestion, liability analysis, evidence collection, injury evaluation, contradiction detection, and settlement packaging. The human BI examiner was engaged at four defined decision points: confirming the intake summary, approving the evidence-gathering plan, validating the damages analysis, and authorising the final settlement offer. The agent did the work; the expert applied judgment at the moments that matter. The result — a 48% reduction from the initial demand inside the 10-day window, with an indemnity saving of \$80,640 — reflects what becomes routinely achievable when experienced examiners are supported by agents operating at L3 Initiates.

## L5 Exceeds: A Simulation of What's Possible

Market research consistently tells us that most carriers are not yet ready to authorise AI agents to handle bodily injury claims at full L5 autonomy — and that is a reasonable position. BI claims carry significant legal, financial, and reputational weight, and the governance frameworks required to oversee fully autonomous settlement decisions in this domain are still maturing. Shift respects and anticipates that readiness curve; it is precisely why the ARISE framework distinguishes L3 from L5 with precision.

We ran a simulation on this same claim, using the same real data, to explore what L5 Exceeds operation would look like — no human checkpoints, the agent acting end-to-end with full authority. In the simulation, the agent did not wait for examiner approval at any stage. It identified the liability reduction opportunity, collected the evidence, made the determination, assembled the settlement package, and submitted the \$87,360 counter-offer autonomously. The third party accepted. Total elapsed time was the same eight days; examiner time was zero.

The simulation is not a product announcement. It is an honest signal about the direction of travel. The technology to operate at L5 on complex BI claims exists today. The question every carrier should be asking is not whether L5 will arrive in injury management — but when their organisation will be ready to meet it, and what governance investments are needed to get there safely. The ARISE framework is designed to help answer exactly that question.

# 5. Implications for insurance leaders

## For Chief Claims Officers and Chief Operating Officers

The SLA framework provides a practical procurement and governance tool. When evaluating AI vendors, require explicit level mapping for each proposed capability. An L3 Initiates agent that presents decisions for one-click approval carries fundamentally different compliance and audit requirements than an L4 Solves agent executing autonomously — and both differ significantly from an L1 Answers chatbot. Clarity at procurement time prevents costly misalignment at deployment time.

Organisations should also establish level-appropriate oversight protocols. L1 Answers and L2 Recommends deployments require minimal new governance. L3 Initiates requires clear exception escalation paths and audit sampling. L4 Solves and L5 Exceeds require portfolio-level monitoring, model-drift detection, and defined thresholds for human escalation — analogous to the operational design domain boundaries that govern SAE L4 vehicle deployments.

## For Chief Risk Officers and Compliance Functions

Regulatory bodies globally are developing frameworks for AI oversight in financial services. The NAIC’s model bulletin on AI, the EU AI Act’s risk-based classification system, and the FCA’s emerging AI guidance all converge on a common principle: the degree of required oversight scales with the

degree of autonomous decision-making authority. The SLA framework operationalises that principle for insurance, giving compliance teams a precise vocabulary to map internal AI deployments to regulatory risk tiers.

Specifically, L4 Solves and L5 Exceeds agents in claims settlement, fraud adjudication, or coverage determination are likely to fall within the “high-risk” AI system category under emerging regulatory frameworks. Organisations that can demonstrate rigorous level classification, documented oversight protocols, and empirical accuracy benchmarks will be better positioned in regulatory examinations than those operating without a formal taxonomy.

## For Technology and InsurTech Ecosystem Participants

Shift Technology invites the broader insurance technology community — carriers, MGAs, third-party administrators, reinsurers, consultants, and industry analysts — to adopt the SLA framework as a common reference. Taxonomy adoption follows network effects: the more broadly a standard is used, the more valuable it becomes. We publish this framework openly and encourage peer review, refinement, and extension as the technology evolves. Our goal is not proprietary advantage through definitional control; it is industry-wide clarity that enables faster, safer, and more accountable AI adoption across the sector.

## 6. Conclusion

The question facing insurance executives is no longer whether to deploy AI agents — it is how to deploy them responsibly, at scale, with appropriate governance and measurable outcomes. That question cannot be answered rigorously without a shared vocabulary for describing what AI agents actually do.

The Shift Levels of Autonomy framework provides that vocabulary. Modeled on the proven taxonomy approach that enabled safe autonomous vehicle development, the SLA framework defines five levels of insurance AI agent capability — the ARISE levels: Answers, Recommends, Initiates, Solves, and Exceeds — with precise definitions, clear examples, and empirical grounding in Shift Technology’s production deployments across the world’s leading carriers.

We believe this framework will become the standard by which insurance AI agents are described, evaluated, procured, and governed. We offer it to the industry not as a marketing document but as a genuine contribution to the infrastructure of responsible AI adoption — because insurance is, above all, a business built on trust. The AI agents that will define its future must be trustworthy, explainable, and precisely understood.

Shift Technology is committed to leading that standard — in the market, in the boardroom, and in the pages of the regulatory guidance that will shape insurance AI for decades to come.

### Selected References

- SAE International. (2021). J3016C: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. SAE International.
- McKinsey & Company. (2023). The State of AI in Insurance: Toward \$1 Trillion in Value. McKinsey Global Institute.
- Deloitte Insights. (2024). 2024 Insurance AI Adoption Survey. Deloitte Center for Financial Services.
- NAIC. (2023). Model Bulletin on the Use of Artificial Intelligence Systems by Insurers. National Association of Insurance Commissioners.
- European Parliament. (2024). EU Artificial Intelligence Act (Regulation EU 2024/1689). Official Journal of the European Union.
- ISO 8373:2021. Robotics — Vocabulary. International Organization for Standardization.

### About Shift Technology

Shift delivers AI agents that transform insurers’ most critical work. By combining deep industry expertise with unrivalled data resources, Shift provides proven results that have earned the trust of hundreds of the world’s leading insurers. Our insurance-grade AI is accurate, explainable and secure, empowering human experts to act with unparalleled speed, complete confidence and a renewed focus on the people they serve.

Learn more at [www.shift-technology.com/es-gb](http://www.shift-technology.com/es-gb)