cXc

A TRADING NAME OF CONXCHANGE PVT LTD

RESEARCH PAPER · DISTRIBUTION SERIES · MAY 2026

# Practical applications of machine learning in insurance distribution

*Beyond the hype, what actually works today.*

v1.0 · Published 18 May 2026 · Classification: Public · Author: cXc Research

*An evidence-led survey of where machine learning and generative AI now create measurable value in carrier-to-distribution and broker-to-customer workflows — and where, despite the volume of vendor noise, they still do not. Written for carriers, MGAs, brokers, and the technology teams that serve them.*

# Executive summary

Distribution is where insurance meets the world. It is also the part of the value chain where the gap between what AI vendors promise and what carriers, MGAs, and brokers have actually deployed remains widest. This paper closes that gap with evidence, not adjectives.

Three observations frame everything that follows.

### One: adoption is high, but maturity is thin

Recent regulator data places UK insurance AI adoption at 95% — the highest of any financial-services sector.[1] Conning's 2025 industry survey shows LLM adoption among US insurers jumped from 18% to 63% in a single year, yet only 7% have scaled AI across the enterprise.[2] The Lloyd's Market Association found that only 14% of London-market firms surveyed have deployed agentic or generative AI in underwriting, while 65% have done so in no part of underwriting or claims at all.[3] Most firms are running experiments, not production systems. The opportunity, and the risk, is in that gap.

### Two: the working use cases are narrower than the marketing suggests

Four categories of application now have measured, repeatable returns in distribution: submission intake and triage, lead and quote conversion modelling, retention and cross-sell, and document understanding. A fifth — algorithmic follow and digital binder placement — is working at scale inside Lloyd's. Beyond that perimeter, results are mixed at best, and the failure cases are instructive: misfiring chatbots, biased pricing, and unsupervised denial automations have produced lawsuits, regulatory consent orders, and headline-grade reversals.

### Three: the regulatory perimeter is hardening, not softening

The EU AI Act classifies life and health insurance pricing as high-risk under Annex III, with deadlines extended by the May 2026 Digital Omnibus agreement to 2 December 2027 for stand-alone systems and 2 August 2028 for product-embedded systems.[4] In the United States, the NAIC Model Bulletin has been adopted by more than 24 jurisdictions, with an NAIC AI Systems Evaluation Tool pilot launched in January 2026 across 12 states.[5] Colorado's bias-testing regime under SB 21-169 now covers auto and health insurance as

well as life. The UK Financial Conduct Authority has signalled that the Consumer Duty is its primary lens for assessing AI in financial services and has no plans for sector-specific AI rules.[6] The carriers and distributors that move first will be the ones whose governance frameworks survive the 2027–2028 procurement cycle. The rest will be replatforming.

**What this paper does**

Each application is presented with a named example of an operator using it at scale, a published number where one exists, and the operational pre-conditions for it to work. Where evidence is mixed or absent, we say so. Where vendor claims do not survive scrutiny, we say so. The goal is a usable map of the territory, not a prediction of the destination.

# Contents

PART I

# The current state

*What is actually deployed, by whom, and at what stage of maturity.*

# 1. From hype to operational discipline

Every two years for the past decade, a McKinsey, Bain, or Capgemini deck has predicted that AI will transform insurance distribution within three years. Every two years for the past decade, the next deck has explained why it has not, and reset the clock. That cycle is breaking now — not because the technology has finally caught up with the rhetoric, but because the rhetoric has finally caught up with what the technology can already do.

The most consequential evidence of that shift is the November 2024 joint survey by the Bank of England and the Financial Conduct Authority. Across 118 UK financial-services firms, 75% reported active use of AI, with a further 10% planning adoption inside three years. Insurance was the highest-adopting sector at 95% — higher than even international banks, at 94%.[1] Foundation models, including large language models, account for 17% of all AI use cases. Critically, only 2% of use cases are fully autonomous; 55% involve some form of automated decision-making, but the typical pattern is augmentation, not replacement.

That distinction matters. The popular narrative of AI in insurance — that intelligent machines will price, sell, and adjudicate without human involvement — is not what the operating practice shows. The dominant pattern is narrower and more useful: AI extracts data, ranks priorities, drafts responses, predicts likelihoods, and flags anomalies, while humans approve, reject, or override. Roughly a third of all AI use cases in the same survey are third-party implementations — up from 17% in the 2022 wave — which signals that insurers are buying capability, not building it from scratch.

## The economic case is now legible

McKinsey's February 2026 analysis estimates that generative AI alone could unlock $50 billion to $70 billion in additional revenue for the global insurance industry, on top of efficiency improvements already in flight.[7] The same firm finds that early AI leaders in insurance are generating roughly six times the total shareholder returns of their AI-laggard peers, and that quoting times in commercial lines have fallen from weeks to days, and in selected sub-segments from days to hours.[7] A Grant Thornton 2025 international survey of

950 insurance executives reports that 52% of insurance leaders already claim AI-enabled revenue growth, and 62% report measurable improvements in decision-making.[8]

These are not forecasts. They are reported operational results across a representative sample of carriers and intermediaries, in published surveys, attached to named individuals at named firms. The character of the AI conversation in insurance has changed accordingly. It is no longer about whether AI works in distribution; it is about which applications work, under what conditions, with what governance, and at what cost.

## Distribution is the part of the chain where the friction is largest

Three structural facts make distribution the most attractive part of the insurance value chain for machine-learning deployment.

First, distribution generates the highest volume of unstructured data: emails, broker submissions, application forms, schedules of values, loss runs, telematics streams, conversational logs. ML and modern document AI thrive on volume.

Second, distribution is where the margin compression is most visible. Bank of America identified more than $15 billion in commissions paid to independent agents in 2025 across just six major US carriers, largely concentrated in low-complexity personal lines and small commercial business.[9] That is exactly the slice of the market where AI-augmented or AI-native distribution is most efficient.

Third, distribution is where new competitors have already crossed the threshold. McKinsey reports that US managing general agent (MGA) direct premiums nearly doubled from $47 billion in 2020 to $97 billion in 2024 — annual growth of about 14%.[7] MGAs typically operate with shorter technology procurement cycles than carriers and have been first movers on AI-enabled triage, pricing, and binder placement. The structural shift that AI is producing in distribution is not abstract; it is showing up in premium-flow data.

**A reasonable working definition**

When this paper refers to "AI in distribution" we mean two distinct families of capability: classical machine learning (supervised, unsupervised, and reinforcement-learning models applied to structured data, used for tasks like lead scoring, lapse prediction, fraud detection, and dynamic pricing), and generative AI (large language models and multimodal foundation models used for document understanding, drafting, conversational interfaces, and increasingly for agentic workflows). The two are not interchangeable. They have different cost structures, different failure modes, and different regulatory treatments. Treating them as one category is the most common source of confused buying decisions in the market today.

# 2. Where AI is actually working: the adoption data

The clearest picture of AI maturity in insurance comes from four credible sources published between November 2024 and May 2026: the Bank of England and FCA Artificial Intelligence and Machine Learning Survey 2024, the Conning 2025 AI and Insurance Technology survey, the April 2025 Lloyd's Market Association seminar on AI use in the London specialty market, and McKinsey's February 2026 analysis of AI investment in insurance. Their findings agree on the broad shape of the curve and diverge productively on its slope.

## Adoption is broad, deep deployment is rare

The headline numbers are easy to summarise:

- In the UK, 75% of financial-services firms surveyed by the Bank of England and FCA are using AI; insurance is the highest-adopting sub-sector at 95%.
- Inside the London specialty market specifically, 14% of firms surveyed by the LMA have deployed agentic or generative AI in underwriting; 65% have done so in no part of underwriting or claims; 47% are experimenting without wide adoption.
- Across all 2024 BoE/FCA respondents, only 2% of AI use cases are fully autonomous; 24% are semi-autonomous; 55% involve some form of automated decision-making.
- In the US, Conning's 2025 industry survey shows LLM adoption among insurers rose from 18% to 63% inside a single calendar year.
- Still on the Conning US data: only 7% of US insurers have scaled AI across the enterprise.

Read together, these data points describe a market that has crossed the experimentation threshold but not yet the operating-system threshold. The dominant pattern is several pilots running in parallel inside any given carrier or distributor, each addressing a discrete workflow problem, with limited integration between them and limited governance coverage. The 2024 BoE/FCA survey found that only 34% of firms claim "complete understanding" of the AI they have deployed; 46% report only "partial understanding",

primarily because they have purchased third-party systems whose internal mechanics they cannot fully inspect.[1]

## Where the deployment is happening

The same surveys are consistent on where AI is actually getting into production. In the London market, the use cases ranked by prevalence are:

- Data extraction from unstructured documents — 74% of LMA survey respondents.
- Submission preparation — 54%.
- Claims triage and automation of simple tasks — approximately one third.
- Policy review and endorsement management — approximately one third.
- Fraud detection — 14% (treated as a primary use case).

This is the operational pattern that this paper will examine in detail. The applications cluster at the points in the distribution workflow where the work is least intellectually distinctive and most volume-bound. They are also, not coincidentally, the points where the case for AI is most defensible to a regulator: augmenting humans on repetitive document work, with the human retaining the binding decision.

## The vendor landscape has consolidated faster than the operator landscape

On the supply side, three patterns are now visible. First, a small number of specialist AI insurance vendors have built scale and reference customers — hyperexponential, Cytora, Send, Shift Technology, Tractable, FRISS, Sixfold, and Artificial Labs being the names that surface most consistently across the survey data and named client deployments. Second, the incumbent core-system vendors — Guidewire, Duck Creek, Sapiens, Insurity — have added AI features to their platform offerings, generally as either embedded predictive analytics or as integrations with the specialists above. Third, hyperscaler models from Anthropic, OpenAI, Google, and Microsoft are increasingly being consumed via Azure OpenAI Service and AWS Bedrock by both vendors and operators directly, with model selection becoming a procurement question in its own right.

The McKinsey investor analysis describes software and data platforms as the fastest-growing category of insurance investment, rising approximately 20% per year over the five

years to mid-2025.[7] In the same period, traditional broker-consolidation deal volume declined by roughly 20%. Capital is rotating into the infrastructure layer of insurance distribution faster than it is rotating into the relationship layer.

# 3. Three claims the evidence supports

Before mapping the use cases in detail, it is worth stating clearly which of the headline claims about AI in distribution are now properly supported by operating evidence. Three are.

## Claim one: ML-driven lead scoring produces measurable conversion lift

Across the credible case-study literature, lead-scoring models built on supervised learning produce a consistent uplift in conversion of two to three times for the top-scored segment and a meaningful efficiency saving by deprioritising the bottom segment. A documented mid-market insurance case showed top-scoring leads converting at 3.5 times the average, with the bottom 20% of leads excluded entirely from agent outreach, producing a 1.5% profit-line improvement in the first months of deployment.[10] Progressive Insurance, the third-largest US auto insurer, has publicly attributed roughly $2 billion in new premium volume to a single ML-driven app-feature improvement.[11] These are not bleeding-edge numbers; they are well-replicated outcomes for a well-understood class of problem.

## Claim two: document AI cuts submission intake time by 60–90%

On submission intake, the evidence is unusually consistent across vendors and operators. Reported reductions in time-per-submission cluster at 60–90% when AI-based intelligent document processing (IDP) is deployed on the dominant ACORD forms (commercial-lines applications 125, 126, 130, 140), with first-pass extraction accuracy of 97–98% on printed fields and 93–96% on handwritten.[12][13] The mechanism is well understood: a commercial-lines underwriter typically spends 30–45% of the working day on data entry and document handling rather than risk evaluation. Automating that work reclaims that capacity without changing the decision boundary.

## Claim three: algorithmic follow placement works at Lloyd's scale

The single most quantitatively transparent example of AI working at scale in insurance distribution is Ki, the algorithmic Lloyd's follow platform launched by Brit in 2020 and separated into Fairfax Group in January 2025. Ki wrote over $1 billion in gross written

premium in 2024 and $1.11 billion in 2025, including more than $200 million of partner capacity from Beazley, QBE, Aspen (Sompo), Travelers, and Tokio Marine Kiln.[14][15] A broker can log into the Ki platform, submit a risk, and receive an algorithmically-priced follow line in as little as ten seconds. That speed is not a marketing claim; it is the operating mechanism that has allowed Ki to scale faster than any new Lloyd's syndicate in the marketplace's recent history.

These three claims share a common structure. They describe narrow, repeatable tasks; they are supported by published operating numbers from named firms; they augment rather than replace human judgement; and they leave the regulatory perimeter intact. They are the cases where AI in distribution is working today.

# 4. Three claims it does not

Three other claims are widely repeated in the market and do not survive contact with the evidence. Operators evaluating AI investments would do well to discount them appropriately.

## Claim one: AI will replace insurance agents and brokers

The shareholder letter and 10-K from Lemonade, the most aggressive AI-native US insurer, are unusually honest about the limits of full automation. As of year-end 2025, Lemonade reports that 96% of first notices of loss are taken without human intervention, and 55% of all claims are end-to-end automated.[16] That second number is the meaningful one: even at the company most architecturally committed to AI-first distribution, 45% of claims still require human handling. The headline 96% FNOL number is a measure of intake capture, not of unsupervised settlement.

Lemonade also continues to operate at a substantial loss — a $1,464.3 million net loss in 2025, with significant catastrophe-driven volatility — and its 2,300-customers-per-employee ratio, while impressive in isolation, has not yet demonstrated sustainable underwriting profitability at scale. The case study of Lemonade is not the case study of agent-replacement. It is the case study of how far conversational AI can take a digital-first insurer before the structural economics of insurance reassert themselves.

Klarna's 2024 decision to replace 700 customer-service staff with AI-powered systems, followed within months by the rehiring of human agents after service-quality and customer-satisfaction metrics collapsed, is the clearest cautionary tale outside insurance proper.[17] The pattern recurs in narrower form in insurance: chatbots that swore at customers, gave incorrect bereavement-fare advice, or made unauthorised promises. The technology is real. The supervision required to deploy it safely in customer-facing distribution is more expensive than the marketing implies.

## Claim two: GenAI will replace underwriters

Recent peer-reviewed research from the underwriting agent-AI literature finds that LLMs hallucinate on 11.3% of commercial-underwriting cases without adversarial-critique architectures in place, and on 3.8% with them.[18] A 3.8% hallucination rate is acceptable in many augmentation use cases. It is not acceptable as a basis for binding commercial coverage without a human underwriter in the loop. Frontier-model benchmarking from late 2025 found that some small frontier model variants have hallucination rates up to 19% on completed underwriting traces.[19] The technology is improving rapidly. It is not yet at the operational threshold for unsupervised commercial-lines underwriting.

This is not a permanent verdict. It is a description of the current state. But the procurement decisions being made in 2026 are committing buyers to specific governance architectures for the next three to five years. Those architectures should be calibrated to what the technology actually does today, not to what it might do by 2030.

## Claim three: AI compliance is principally an IT problem

The regulatory section of this paper sets out in detail why this third claim is wrong. The EU AI Act, the NAIC Model Bulletin, Colorado's SB 21-169, and the UK's principles-based Consumer Duty framework all treat AI governance as an actuarial, legal, and senior-management responsibility rather than primarily a technical one. The most common procurement failure currently observable in the market is treating AI implementation as a CTO project and then discovering, sometimes through a market-conduct exam, that the documentation, bias-testing, and human-oversight architecture required to defend the system was not built into procurement. The cost of retrofitting that governance after the fact is significantly higher than the cost of building it correctly the first time.

> **The disciplined posture**
>
> Treat AI in distribution as an augmentation problem with hard governance constraints, not as an automation problem with soft engineering constraints. The carriers and distributors that have followed this posture so far have measurably better

outcomes than those that have not. The evidence on this is consistent across surveys, peer-reviewed research, and published case studies.

PART II

# Practical applications in distribution

*Eight workflow areas, each with operating evidence, named operators, and a candid view of pre-conditions.*

# 5. Lead acquisition and intent prediction

Lead scoring is the oldest and most thoroughly evidenced AI use case in distribution. It is also, for that reason, the application where the gap between what is possible and what most distributors actually do is largest. A surprising share of agencies and direct distributors still allocate outreach by first-in, first-out queueing or by simple firmographic filters, despite a decade of evidence that supervised-learning models materially outperform these approaches on conversion rate.

## What the models actually do

Modern lead-scoring models for insurance use a combination of inputs: traditional firmographics (industry SIC code, revenue band, headcount, geography); behavioural data (website pages viewed, content downloaded, time-to-response); third-party signals (recent funding rounds, executive movements, regulatory filings, hiring patterns); and where available, prior policy data (claims history, premium paid, lapse pattern). The output is a propensity-to-convert score, typically with explanatory feature importances attached for compliance and producer trust.

The published case-study numbers are remarkably consistent. The Intelliarts predictive lead-scoring deployment for a US-based insurer cut roughly 6% of unproductive leads, produced a 3.5x conversion uplift for the top-scoring segment, and improved overall profit by 1.5% within months of deployment.[10] The Datagrid and Symphonize implementations report similar conversion-lift ranges across commercial-lines agencies. The technology behind these results is not exotic — gradient-boosted decision trees (XGBoost, LightGBM, CatBoost) and increasingly transformer-based architectures for the unstructured-data inputs.

## Where it works, where it fails

Lead scoring works best where the distributor has three things: a sufficient volume of historical conversion data (typically at least 5,000 closed-loop outcomes); a clean attribution mechanism that ties marketing inputs to underwriting outcomes; and a producer team willing to act on the scores. The third condition is the one most often

underestimated. A score that producers do not trust is a score that does not change behaviour. Most credible vendor case studies bury this prerequisite in the implementation phase.

It fails, or underperforms, in three contexts. First, where the line of business is so high-touch and bespoke (London market specialty, large-account commercial) that the conversion event itself is too rare to support model training. Second, where the firmographic signal is dominated by a relationship dimension the model cannot see (a producer's golf-club network is a real conversion driver and one no model can score). Third, where regulatory constraints on input features — Colorado's SB 21-169 testing requirements, for instance — eliminate variables the model relies on for predictive power.

## Producer-side and broker-side variants

Two structural variants of lead scoring are worth distinguishing. The producer-side variant scores prospect-to-policy conversion: it ranks new prospects for a single producer or agency. The broker-side variant scores broker-to-carrier appetite match: when a broker is preparing to market a risk, it identifies the carriers most likely to bind that specific risk type at competitive terms. The second variant is the more sophisticated and the harder to deploy, because it requires multi-party data sharing. The most credible recent deployment is hyperexponential's Triage product, launched in September 2025, which is positioned exactly at this broker-to-carrier appetite-matching layer.

Hyperexponential's platform processes $45 billion in premium annually across customers including Markel and Sompo, and its underlying thesis is that the carrier appetite-match problem and the underwriter prioritisation problem are best solved with the same data infrastructure.[20] That is a defensible architectural argument and increasingly the dominant pattern in commercial-lines distribution technology.

**Pre-conditions to deploy**

A working lead-scoring deployment in insurance distribution typically requires: a CRM with 18–36 months of clean attribution data; clearly defined conversion events;

producer-incentive alignment with the scoring outputs; a bias-testing protocol if the line of business touches consumer or small-commercial classes in Colorado or any NAIC Model Bulletin jurisdiction; and a governance framework that can document why each input feature is used. The build phase is typically 12–16 weeks; the change-management phase is typically longer.

# 6. Submission intake, classification, and triage

If lead scoring is the oldest application of ML in distribution, submission intake and triage is the application where modern generative AI has produced the largest step-change in operational impact. It is also the application where the strongest current consensus exists between vendors, operators, and regulators about what works.

## The problem in operational terms

A commercial-lines underwriter receives on average 40 to 60 submissions per week. Each submission is typically a bundle of an ACORD 125 (commercial insurance application), one or more line-specific supplements (126 for general liability, 130 for workers' compensation, 140 for property), three years of loss runs from prior carriers, financial statements, supplemental questionnaires, and a broker cover narrative. The documents arrive in different formats from different brokers, scanned at varying quality, often with handwritten margin notes. The underwriter's first task before any risk evaluation begins is to read every page, find the relevant data points, and key them into a workbench or rating engine.

Industry analyses converge on the same headline: that intake and data-handling work consumes 30 to 45% of an underwriter's working day before any risk evaluation begins. Carriers quote, by their own data, only about half of the submissions they receive. The bottleneck is not underwriting expertise; it is the speed at which broker submissions can be ingested, normalised, and routed to the underwriter best equipped to write them.

## What the AI now does

Modern intake and triage systems combine three capabilities: document classification (identifying that a given attachment is an ACORD 125 versus a loss run versus a financial statement); extraction (pulling structured fields out of the documents, including from handwritten annotations); and triage (scoring the submission against the carrier's appetite, capacity, and capacity remaining, and routing it to the appropriate underwriter).

The reported operating results are unusually consistent. SortSpoke reports 60–90% reduction in intake time per submission across its agency customer base.[13] Lido reports

97–98% accuracy on printed ACORD fields and 93–96% on handwritten.[12] Underwriting Software's deployment reports a 25-minute-to-2-minute reduction in submission triage time. A 1,200-submission renewal queue that previously required six staff on overtime can now be processed without temporary headcount.[21]

Hyperexponential's Triage product, launched in September 2025, integrates the intake function with carrier pricing and rating engines so that underwriters can issue indicative pricing within minutes of submission receipt, without any additional data entry.[20] The integration is the point. A standalone IDP product saves intake time. An integrated intake-to-pricing pipeline produces a structural change in the time-to-quote, which is the variable that determines broker placement decisions.

## Lloyd's and the London market case

In April 2025, the Lloyd's Market Association surveyed 81 firms, including 45 managing agents representing the bulk of the market's £56.2 billion 2025 stamp capacity.[3] The single most prevalent AI use case was data extraction from unstructured documents, used by 74% of respondents. 54% used AI for submission preparation.[3] The leading vendors named in the same market are Cytora, Sixfold, and Artificial Labs' Ava — each focused on the submission-intake and triage layer.

Cytora's published case studies include Zurich, which the vendor reports has scaled agentic AI across five countries in 90 days, and Markel, which has reported productivity gains of over 100% on the workflows the platform handles.[22] The Markel number, in particular, is unusual in being a doubling of productivity rather than a percentage-point improvement, and it points to where commercial-lines distribution is now headed: the same underwriter writing twice the business, with the AI handling the intake and triage that previously absorbed half the working day.

## Why this category is structurally important

Submission intake is the use case where AI most directly addresses the distribution capacity constraint. A commercial-lines underwriter is the most expensive seat in distribution. The cost of generating one more skilled underwriter is high, and the supply of

senior commercial underwriters is constrained. AI-augmented intake creates effective underwriting capacity from existing staff. That is more valuable, particularly in a hardening market, than any of the customer-facing AI applications that get more attention in the trade press.

| Vendor | Geography | Reported result | Source |
|---|---|---|---|
| hyperexponential Triage | Global, UK+US focus | Indicative pricing within minutes; $45B premium processed annually | Vendor (2025) |
| Cytora | Global, Lloyd's heritage | Zurich: 5 countries in 90 days; Markel: 100%+ productivity uplift | Vendor (2026) |
| Sixfold | US commercial lines | Submission scoring against appetite; named LMA leader | LMA survey (2025) |
| Artificial Labs (Ava) | Lloyd's market | Submission triage and Smart Follow auto-bind | LMA survey (2025) |
| SortSpoke | North America, agency | 60–90% reduction in intake time per submission | Vendor (2026) |
| Lido | Mid-market, US-led | 97–98% first-pass accuracy on printed ACORD fields | Vendor (2026) |

# 7. Quote-and-bind: from days to seconds

The quote-and-bind time has historically been the variable that determines competitive position in personal lines and increasingly in small commercial lines. The distributor that returns a bindable quote first, with adequate accuracy, wins disproportionately. AI has now compressed this variable to a degree that changes the structural economics of distribution.

## Personal lines: the digital-first benchmark

Lemonade reports that more than 90% of its policies are sold through conversational AI without any human intervention from quote to bind. The conversational onboarding agent, branded Maya, asks 13 questions in the user-facing flow but collects more than 1,600 underlying data points to feed risk-scoring and pricing models in real time.[16] The end-to-end policy-bind process operates inside 90 seconds in the typical case. This is not a research result; it is the operating practice of a publicly listed insurer at the scale of approximately three million customers.

Lemonade's pet insurance line grew 55% year-over-year in 2025 (from $283 million to $439 million of in-force premium), a growth rate that outpaces incumbent competitors by as much as 350%.[23] The conversational quote experience is the proximate cause; the data flywheel it produces is the structural reason the lead persists. Each conversation captures behavioural and intent data that improves segmentation and pricing, which in turn produces a more competitive product. That feedback loop is now visible in market-share data, not just in product theory.

## Commercial lines: from weeks to days

Commercial-lines quote-and-bind is a different problem. The risk is more bespoke, the data inputs are more varied, and the bind decision typically requires human approval for any account above a threshold size. The AI compression here is correspondingly less dramatic but still operationally consequential.

McKinsey's 2026 analysis observes that quoting times for commercial lines have fallen from weeks to days at carriers that have deployed integrated AI-pricing platforms, and in selected

sub-segments from days to hours.[7] Hyperexponential's customers — including Markel, Sompo, and a growing US admitted-lines client base — represent the operational evidence base for these claims.[20]

## Embedded distribution and API-first models

A related and increasingly important variant of quote-and-bind compression is embedded insurance, where coverage is offered at the point of purchase of another product or service. Mordor Intelligence values the global embedded insurance market at $13.88 billion in 2025, projected to reach $68.12 billion by 2031 at a 30.4% compound annual growth rate.[24] Approximately 65% of embedded insurance providers now use AI for underwriting accuracy and personalised policy recommendations.[25] Zurich's April 2025 partnership with Ominimo launched AI-tuned motor insurance across three countries in eight weeks — a deployment timeline that would have been impossible without modern API-first integration and AI-driven underwriting infrastructure.[26]

## What carriers and MGAs are actually building

The procurement pattern that has now consolidated is a three-layer architecture: an intake and document-AI layer (hyperexponential, Cytora, Send, or in-house equivalents); a rating-and-pricing engine layer (hyperexponential, Earnix, Akur8, or carrier-proprietary); and a binding-and-policy-administration layer (Guidewire, Duck Creek, Sapiens, Insurity, or modern cloud-native equivalents). The AI layer is currently most mature at the intake end and the pricing end. The bind-and-administer layer is where AI penetration is shallowest, in part because policy administration systems carry the highest regulatory inspection burden.

**What the operating numbers suggest**

Personal lines: 60–90 seconds end-to-end is the new benchmark for digital-first quote-and-bind. Anyone slower is losing share to digital natives.

Small commercial (under $25k premium): hours to days is the new benchmark; weeks is no longer competitive.

Mid-market and specialty: hours-to-quote is achievable on appetite-aligned business; days is the operating reality on most accounts; bespoke specialty risks remain a relationship-led, multi-week process where AI augmentation is most valuable in submission triage and pricing data preparation, not in displacing the underwriter.

# 8. Producer and agency performance analytics

Distribution at every scale generates performance data that should, in principle, be a goldmine for ML. In practice, this is the application area where the technology has been most slowly adopted, in large part because the politics of producer compensation make any predictive intervention sensitive.

## The metrics that matter

Three families of producer metrics are now routinely modelled with supervised learning. The first is quote-to-bind ratio (the percentage of submitted quotes that convert to bound policies, both at agency and individual producer level), and the closely-related producer hit ratio. The second is account-rounding propensity (the likelihood that a single-line policyholder will accept a multi-line bundle when offered). The third is retention forecasting at producer level, which combines the previous two with renewal-cycle data.

The major agency management system vendors — Salesforce, EZLynx, Applied Epic, AMS360 — now ship predictive-analytics dashboards built on these metrics.[27][28] In the US, the Insurity Predict offering, built on the Valen Data Consortium of $109 billion in premium across all standard US P&C lines, claims 3–10% better loss ratios at deploying customers through predictive variables fed into underwriting and pricing decisions.[29]

## Where it actually changes behaviour

The applications that materially change distribution behaviour, in the published evidence, are the following four:

- Account-rounding propensity scoring. A model identifies the top 5–10% of single-line policyholders most likely to accept a bundle if approached, and routes them to producers in priority. The implementation cost is low; the conversion uplift on the targeted segment is typically large; the political resistance from producers is minimal because the work is additive rather than reallocative.
- At-risk renewal prediction. A model identifies the top 10–15% of renewals most likely to lapse without intervention, allowing producers to allocate retention effort by risk-adjusted expected value rather than by alphabetical order or last-month-of-

policy timing. This is among the highest-ROI distribution AI applications when measured by retention dollars per producer hour.

- Carrier-appetite matching for placement. On the broker side, a model identifies the carriers most likely to bind a given risk at competitive terms, reducing the number of unsuccessful submissions and producer time spent on placements that would not have bound.
- New-producer ramp prediction. Less common but high-value: a model predicts which new producers in their first 12 months will hit retention thresholds, identifying coaching candidates early.

## Where it generates resistance

Performance prediction models that flag specific producers as low-performing typically meet active resistance, and reasonably so. The producer-as-individual is not a useful unit of prediction for most agencies, because conversion is so dependent on book composition, geography, and incoming-lead quality that comparable benchmarking is methodologically difficult. The most successful deployments treat producer performance analytics as a coaching tool rather than a ranking tool, and put producer trust ahead of model accuracy in the implementation sequence.

## Commission and reconciliation analytics

A related but separately deployable application is AI-augmented commission management. The commission-reconciliation problem — matching incoming carrier commission statements against expected schedules, identifying discrepancies, and routing corrections — is a high-volume, low-judgement task where supervised learning and modern document AI produce immediate operating returns. The administrative cost of commission management at a typical mid-sized broker is non-trivial, and the audit-trail benefit of AI-assisted reconciliation extends beyond the operating saving.

> **A note on the political constraint**

Producer performance analytics in distribution is the AI application where the implementation politics most often determine the outcome. Models that are framed as helping producers — improving the quality of leads routed to them, prioritising their renewal effort, identifying high-propensity cross-sell targets — succeed. Models that are framed as evaluating producers tend to be quietly disused, regardless of their statistical merit. The most experienced operators sequence the deployment carefully.

# 9. Retention, lapse risk, and cross-sell modelling

Insurance distribution economics are heavily lifetime-value driven. A policy held five years is worth materially more than the first-year commission suggests. Lapse prediction and retention modelling have been a stable AI use case for a decade, and the technology is now well-developed, but the application of those models to producer behaviour remains uneven.

## What works in the model

Lapse-prediction models in insurance typically combine policy-level data (payment history, claims history, policy-change frequency, premium increases) with policyholder behavioural data (interaction with the carrier's app, response to marketing, channel mix) and external data (economic conditions, employment status proxies, life-event signals). The dominant technical approach is gradient-boosted decision trees (XGBoost, LightGBM, CatBoost), with accuracy rates of 85–95% in well-instrumented portfolios.[30]

Reported research from the French insurance market using the Grabit model — a hybrid of supervised learning with the Tobit statistical approach — has produced churn-prediction accuracies superior to either traditional regression or XGBoost alone in published peer-reviewed work.[31] In the life-insurance market, EIOPA's Quantitative Impact Study QIS5 identifies lapse risk as the most significant life-underwriting risk, which puts a regulatory premium on accurate prediction.[32]

## What works in the operating use

The technical accuracy of these models is the easy part. The operational deployment patterns that have produced measurable retention uplift in published case studies share three characteristics.

- Intervention is calibrated to predicted lapse probability and policy value. A high-probability, high-value lapse warrants a personal outreach from a senior producer or account manager. A low-probability lapse warrants no intervention. A high-probability, low-value lapse may warrant only an automated email.

- Intervention timing is calibrated to the lapse window. A policy at high lapse risk 60 days before renewal is a different intervention from one at high lapse risk 5 days before renewal.
- Intervention content addresses the underlying cause where the model can identify it. A lapse driven by premium increase warrants a value-justification conversation; a lapse driven by service issue warrants a complaint-resolution outreach.

The carriers and distributors that get measurable retention uplift from lapse prediction do all three of these things. The ones that do not, often because the model output is treated as a uniform alert list, see little operational return on the modelling investment.

## Cross-sell modelling: a quieter success

Cross-sell modelling — predicting which existing policyholders will accept which additional products — has been less written about than lapse prediction but is arguably the higher-value application for distributors. The core technique is a propensity model trained on prior cross-sell outcomes, with the additional product as the target variable. The output is a propensity score and a recommended-product ranking per policyholder.

Lemonade's published 2025 data shows that bundling and cross-sell motion delivered a meaningful share of its overall in-force premium growth to $1.24 billion (a 31% year-over-year increase).[16] The mechanism is that the conversational AI already knows the customer, and the cross-sell is offered naturally inside the next service interaction rather than as a separate marketing campaign.[23] For traditional distributors, the equivalent mechanism is producer-routed: the model identifies the prospect, the producer makes the conversation, the AI handles the back-office work.

## A note on regulated retention practices

UK and EU regulators have signalled active interest in retention pricing and the use of AI in that context. The FCA has been clear that the Consumer Duty applies to renewal pricing decisions, and that AI-driven retention pricing must satisfy the fair-value obligation.[33] In practice, this means retention-uplift modelling must be paired with a documented fair-value framework that demonstrates the resulting prices remain proportionate to the benefit delivered. Distributors that operate retention pricing without that framework are exposed.

# 10. Conversational interfaces in distribution

The most visible category of AI in insurance distribution is the conversational interface — chatbots, voice agents, and AI-augmented customer-service interfaces. It is also the category where the gap between vendor demos and production reality is widest, and where the deployment failures have produced the most reputational and regulatory cost.

## Where it works

Conversational AI works well, and is now standard practice, in three contexts.

- Quote intake for low-complexity personal lines. The Lemonade Maya pattern (13 questions, 1,600 data points, 90-second policy bind) is the dominant template across the US digital-first carrier landscape. US auto insurer GEICO's Kate handles over 100 million interactions annually in a similar but more constrained role.
- First notice of loss in personal lines. The Lemonade AI Jim pattern (96% of FNOL taken without human intervention) is well-established in the US and now mirrored in less-publicised form across most major carriers globally.
- Policy-management self-service. US carrier Allstate's Virtual Assistant Suite handles 60% of routine inquiries autonomously, escalating only the cases that genuinely require human handling. The cost saving on routine interactions is significant; the customer-experience trade-off is now usually neutral or favourable, where the deployment is well-engineered.

## Where it does not, yet

Conversational AI underperforms in three contexts that matter to distribution.

- Complex claims handling. Beyond a relatively simple FNOL capture, the conversational AI handoff to human handler is fragile; the customer experience is often worse than a fully-human channel; and the regulatory exposure on automated denial decisions is now significant.
- Producer-facing tools. Despite years of vendor claims, the producer-facing AI assistant that actively helps an experienced producer write better business has been hard to deploy. Salesforce's Agentforce, the Insurity producer-AI tooling, and

similar offerings have improved on prior generations but remain a complement to, rather than a replacement for, experienced producer judgement.

- Cross-channel handoffs. The conversation that begins on a website chatbot, escalates to a voice channel, and then returns to email is the case where most conversational AI deployments lose context. The customer experience suffers; the regulator interest in record-keeping across the channels makes the engineering harder.

## The Lemonade lesson, properly understood

Lemonade is the most-cited case study of conversational AI in insurance, and the most often misread. The headline numbers (96% FNOL automation, 55% end-to-end claims automation, 2,300 customers per employee, 90% of policies sold via bot) are real and impressive.[16][23] They are also accompanied by less-discussed realities: substantial annual net losses (a $1,464.3 million net loss in 2025); a 2021 lawsuit alleging biometric data mishandling that materially damaged customer trust; a heavy dependence on reinsurance (55% of the book ceded over current treaty periods); and concentration risk in catastrophe-exposed lines.[16]

The point is not that the Lemonade architecture is wrong. It is that the architecture is more expensive to operate, and more dependent on continued capital inflow and reinsurance access, than the headline-number narrative suggests. Carriers and distributors evaluating conversational AI should plan for the full cost of operating it, including the supervision and regulatory overhead, not just the efficiency saving on routine interactions.

## The 2026 watermarking requirement

Under the May 2026 EU AI Act Digital Omnibus agreement, watermarking and provenance-labelling obligations for AI-generated content apply from 2 December 2026 — earlier than the original 2 August 2026 date in some respects.[4] Any customer-facing chatbot, voice agent, or synthetic-content tool that produces text, images, audio, or video to EU customers will need to support technical traceability. Distributors operating cross-border should treat this as an immediate procurement criterion rather than a future planning item.

# 11. Algorithmic follow and capacity orchestration

Of all the practical applications of AI in distribution, algorithmic follow placement is the one where the operating evidence is least equivocal. The technology works, the numbers are public, and the structural shift in the Lloyd's market it is producing is now visible in premium-flow data.

## What it is, in operational terms

In the London market and increasingly in other specialty markets, large risks are placed across multiple syndicates or carriers: a lead underwriter takes the first share and sets the terms, and follow underwriters take subsequent shares of the same risk at the same terms. The follow market has historically been a relationship-led, manually-administered process, with brokers physically walking from box to box, or its electronic equivalent, to fill out the placement. The algorithmic follow model replaces the manual scouring of the follow market with an algorithm that prices the broker's submission, decides how much capacity to offer, and commits the line — all in seconds.

## Ki: the canonical case

Ki was incubated by Brit in 2020, began underwriting as Syndicate 1618 on 1 January 2021, and grew rapidly: $1 billion in gross written premium in 2024, $1.11 billion in 2025.[14][15] On 1 January 2025, Ki separated from Brit to become a standalone subsidiary within Fairfax Group, retaining Asta Managing Agency as its managing agent.[34] Capacity partners now include Beazley, QBE, Aspen (Sompo Group), Travelers, and most recently Tokio Marine Kiln, bringing the platform to six Lloyd's syndicates.[15]

The operating mechanism: a broker logs into the Ki platform, enters details about the risk being placed, and the algorithm returns a follow line in as little as 10 seconds. The algorithm was developed jointly with Google Cloud and University College London's Computer Science department. It underwrites the risk against the lead's terms, decides how much capacity to offer based on the carrier's appetite and the platform's portfolio composition, and binds the line. Brokers can complete placement of a complex specialty risk on the Ki platform in minutes.[35]

## Why this matters for distribution beyond Lloyd's

The Ki model is not specific to Lloyd's. The underlying pattern — an algorithm prices and binds capacity for repeated, well-characterised risk types against pre-defined appetite, with human underwriters retained for the lead role and for risks outside the algorithm's training distribution — is generalisable. Variants of the model are now visible in:

- Reinsurance: algorithmic capacity allocation for property catastrophe layers.
- US E&S markets: AI-augmented binding authorities for select sub-segments.
- Embedded insurance: API-based capacity provisioning at the point of sale for retail platforms.
- Personal lines: dynamic capacity allocation across geographic and product cells.

The economic case is the same in each variant: the algorithm allows carriers to participate in risk types and volumes that would otherwise exceed available skilled-underwriter capacity, while retaining underwriting discipline because the algorithm itself is constrained by the carrier-defined appetite parameters.

## The governance architecture

Algorithmic follow operates inside a governance architecture that is now well-developed, partly because Lloyd's itself has set expectations through its Future at Lloyd's and Blueprint Two programmes. The core elements:

- Carrier-defined appetite constraints expressed as algorithmic parameters.
- Continuous portfolio-monitoring against accumulations, rate-adequacy, and class-of-business mix.
- Pre-defined human-escalation thresholds for risks outside the algorithm's authority.
- Audit trail of every algorithmic decision, with full traceability to the input data.
- Periodic actuarial review and model recalibration with documented sign-off.

This is the architecture that carriers in other markets — US E&S, EU non-life, embedded distribution — are now studying as the template for algorithmic capacity provisioning. The Lloyd's experience is unusually well-documented, partly because the market has thirty

syndicates and a coordinating Corporation that publishes operating data and survey results. The lessons travel.

**The structural point**

Algorithmic follow is the application of AI in distribution where the value created is not efficiency or speed but capacity itself. The carrier can write a class of business at a volume and granularity it could not otherwise sustain, because the algorithm extends skilled underwriting attention to a larger surface area than human underwriters alone could cover. This is the most strategically important AI use case in distribution, even though it gets less marketing attention than chatbots.

# 12. Fraud and sanctions screening at the distribution layer

Fraud and sanctions screening are not the same thing, but in operating practice in distribution they share an architecture: incoming flows of customer, broker, and claimant data are screened against external reference data and internal pattern libraries, with anomalies routed for human review. AI has reshaped both, in distinctive ways.

## Fraud detection in claims and applications

Shift Technology, the most-deployed dedicated fraud-AI vendor in insurance globally, has reported customer outcomes including: a top-five US insurer identifying fraud networks through shared-data approaches; real-time detection that stopped a $91,000 fraudulent claim; Shift Subrogation delivering 4x ROI in year one for one carrier with a 60% acceptance rate.[36] Tokio Marine Kiln, Covéa, and AXA Switzerland are among Shift's named European deployers.

Shift launched Shift Claims in September 2025, an agentic-AI claims-management platform that integrates fraud detection with claims assessment and triage.[37] Early-adopter results reported by the vendor: 3% lower claims losses, 30% faster claims handling, 60% overall automation rate, and greater than 99% accuracy in claims assessment. AXA Switzerland is the named early adopter.[38]

Lemonade's Forensic Graph fraud-detection system, applied across the customer journey rather than at the claims step, has documented millions of dollars in prevented losses through machine-learning identification of complex multivariate links across signals that humans could not feasibly aggregate.[16] The underlying technique — graph-based machine learning on entity relationships — is increasingly the standard architecture for distribution-layer fraud detection.

## Sanctions screening: an underrated AI use case

Sanctions screening is, in volume terms, the largest single AI-relevant compliance workflow in insurance distribution. Every new policy, every claim, every reinsurance cession, every

broker payment must be screened against sanctions lists maintained by OFAC, the UK Office of Financial Sanctions Implementation, the EU consolidated list, and an expanding constellation of national regimes. The volume of false positives produced by traditional name-matching algorithms is operationally significant: typical false-positive rates of 95% or higher are reported across the financial-services sanctions-screening space.

Modern AI-augmented screening produces material reductions in false-positive rates through three mechanisms: better entity-resolution (correctly identifying whether two name-variants are the same entity); contextual scoring (using transactional context to deprioritise low-risk matches); and continuous re-screening (where existing customers are continuously re-screened against updated lists). The combined effect is a substantial operational saving in compliance-team time, paired with improved screening quality.

Sanctions-screening AI sits inside a regulatory architecture that is particularly demanding: under both EU and US regimes, the responsibility for sanctions compliance cannot be outsourced, and the documentation of how every decision was made must be auditable. This is the application area where the audit-trail and explainability requirements of insurance distribution AI are tightest, and where the procurement bar for AI vendors is correspondingly highest.

## Distribution-layer fraud: an underexamined category

Most published AI-fraud work in insurance focuses on claims fraud. A growing but underexamined category is distribution-layer fraud: false applications, identity manipulation in onboarding, broker-side fraud, and producer-side fraud. The data infrastructure required to detect distribution-layer fraud — cross-carrier entity-resolution, behavioural analysis of application patterns, anomaly detection at the policy-bind step — has historically been underdeveloped because no single carrier sees enough of the market to train robust models. Industry-shared data approaches, including those underwritten by NICB in the US and IFB in the UK, are now combining with AI to close that gap, and distribution platforms that orchestrate the data layer are positioned to benefit disproportionately from the resulting models.

| Use case | Mature? | Operating evidence |
|---|---|---|
| Claims fraud (general) | Yes | Shift Technology 4x ROI; multiple Tier-1 carriers; FRISS deployments across European market |
| Document fraud (image reuse) | Yes | Shift specific case study: detected reused photo fraud across multiple claims |
| Identity manipulation in onboarding | Maturing | Banking-derived techniques now applied; graph-ML standard architecture |
| Distribution-layer fraud (producer/broker) | Emerging | Cross-carrier data sharing required; industry-shared databases are the enabling layer |
| Sanctions screening (false-positive reduction) | Yes | Mature in banking; in-flight in insurance; OpenSanctions and incumbent vendors both AI-augmenting |
| AML transaction monitoring (intermediary payments) | Yes | Banking-grade tooling now deployed across MGA and reinsurance flows |

PART III

# Where ML and GenAI struggle

*The failure modes that matter and what they mean for procurement.*

# 13. The failure cases worth learning from

The published failure cases in insurance AI cluster into four categories. Each is instructive, and each carries an operating lesson that is not always drawn from the case in trade press coverage.

## Failure mode one: automated denial without proper human oversight

Health insurance utilisation review provides the best-documented failure mode for insurance AI. Recent research in Health Affairs describes the structural risks of AI use in prior authorisation and claims processes: the role of the human-in-the-loop is often weaker than vendors claim; users' lack of clinical expertise heightens the risk that AI hallucinations go uncorrected; the high volume, time pressure, and repetitive nature of utilisation review compounds the risk of automation bias.[39] The pattern generalises: AI applied to denial-style decisions at scale, with weak human review, produces both higher denial rates and higher subsequent litigation exposure.

This is the failure mode that has produced the most regulatory consent orders, class actions, and legislative responses to date. It is also the failure mode that the EU AI Act, the NAIC Model Bulletin, and Colorado's SB 21-169 are most directly addressing.

## Failure mode two: chatbot hallucination and unauthorised commitment

The Air Canada bereavement-fare case — where an airline chatbot promised a customer a refund that did not exist, and a tribunal subsequently held the airline accountable for the chatbot's statements — is the canonical case. Similar failures have surfaced in insurance distribution: chatbots giving misleading policy advice, accusing high-profile individuals of criminal damage, swearing at customers, and inappropriately providing financial guidance.[40] Each case has produced reputational damage; some have produced regulatory or contractual liability.

Armilla's AI insurance policy, underwritten by several Lloyd's insurers, was launched in 2025 specifically to cover legal claims arising from chatbot errors and hallucinations, including in cases of model drift or model collapse.[40] The availability of that insurance is

itself a market signal: this failure mode is now sufficiently well-characterised that it can be priced and underwritten as a discrete risk.

## Failure mode three: bias and disparate impact in pricing

The Lemonade biometric-data lawsuit of 2021 — alleging that the company stored facial-recognition data and used non-verbal cues in fraud detection in ways that risked racial and demographic bias — produced reputational damage that the company is still managing in 2026.[41] The lawsuit was settled, but the broader concern about AI pricing models producing disparate impact, even where they do not use protected characteristics directly, has crystallised into specific regulatory regimes.

Colorado's SB 21-169 and Regulation 10-1-1 require quantitative testing for disparate impact in insurance pricing — even where the insurer does not believe its system discriminates.[42] Effective 15 October 2025, the scope expanded from life-only to include private passenger automobile and health benefit plans.[42] The 2025 NAIC AI Bulletin, anticipated to be released alongside the 2023 bulletin's continued state adoption, is expected to strengthen bias-testing expectations across the broader US market.

## Failure mode four: full-replacement automation reversal

Klarna's decision to replace 700 customer-service staff with AI-powered systems in 2024 — followed within months by the public acknowledgement that service quality had collapsed and customer satisfaction had fallen, and the rehiring of human agents — is the dominant cautionary example outside insurance.[17] The pattern recurs in narrower form across multiple consumer-facing service operations: aggressive automation rollouts driven by cost-reduction targets, followed by reversal once the deferred operational cost (escalations, complaints, regulatory attention) becomes visible.

This pattern is unusually instructive for insurance distribution because the cost of automation failure in insurance is structurally higher than in retail: the regulatory exposure is greater, the customer relationship is longer-cycle, and the trust capital is more expensive to rebuild. Insurance distributors that automate aggressively without holding back staffing

capacity for escalations are buying a position they may need to reverse, at a cost that exceeds the saving they were chasing.

# 14. Hallucination, bias, and the human-in-the-loop question

Three operating risks of AI in distribution deserve separate treatment because they apply across the use-case categories rather than being specific to any one of them.

## Hallucination

LLM hallucination — the production of plausible-sounding but factually incorrect output — is now well-quantified in the insurance underwriting context. Recent peer-reviewed work on agentic AI for commercial insurance underwriting found base-rate hallucination of 11.3% without adversarial-critique architectures, falling to 3.8% with them.[18] Frontier-model benchmarking on underwriting tasks has shown some small frontier-model variants producing hallucinations on up to 19% of completed traces; pass@k results showed a 20% drop in answer correctness when models were tested for reliability rather than peak accuracy.[19]

Operating implications: any LLM-based application in distribution that touches a binding decision, a pricing decision, or a denial decision must have an architecture that produces explicit hallucination-mitigation. Adversarial self-critique, retrieval-augmented generation with verifiable sources, structured-output enforcement, and human-in-the-loop verification are all viable techniques. None of them are free. Procurement decisions that do not include explicit hallucination-mitigation requirements are accepting a risk-tolerance that the regulators will not.

## Bias

Bias risk in distribution AI is now actively regulated in at least three jurisdictions: Colorado (via SB 21-169 and Regulation 10-1-1), the EU (via the AI Act's Annex III high-risk classification for life and health insurance risk and pricing assessment), and the UK (via the FCA's Consumer Duty fair-value obligation applied to AI-driven pricing decisions).[4][42][43] The standard for bias-testing varies across these regimes — Colorado is more quantitative, the EU more procedural, the UK more outcomes-focused — but the underlying expectation is the same: AI systems used in pricing or denial decisions must be tested for disparate

impact, the testing must be documented, and the documentation must be made available to regulators on request.

The most common procurement failure in this area is treating bias-testing as a vendor-provided certification rather than as an in-house responsibility. Both the NAIC Model Bulletin and Colorado's regulations are clear that the responsibility for AI bias-testing remains with the insurer, regardless of whether the model was built in-house or supplied by a third party. The model card or fairness certificate provided by a vendor is a useful input; it is not a substitute for the operator's own bias-testing programme.

## Human-in-the-loop: the architecture question

Across the BoE/FCA survey responses, 55% of AI use cases have some automated decision-making, but only 2% are fully autonomous. The dominant operating pattern is some form of human-in-the-loop architecture. The quality of that architecture varies dramatically across deployments. Three considerations distinguish robust from fragile architectures.

- Trigger sensitivity. The human reviewer must be triggered by genuinely informative signals — confidence thresholds, anomaly detection, regulatory exception flags — not by a uniform sampling rate. A 5% random review is theatre; a 100% review of low-confidence outputs is governance.
- Reviewer authority. The human reviewer must have real authority to override the AI, in practice as well as in policy. Architectures that nominally include a human reviewer but routinely default to the AI output are not in fact human-in-the-loop architectures; they are automation with documentation.
- Audit trail integrity. Every human override, and every human approval, must be logged with timestamp, reviewer identity, and rationale. This is the audit-trail requirement that produces robust governance and that satisfies regulator inquiries; it is also the operating cost that is most often underestimated in initial procurement.

The operational cost of doing human-in-the-loop architecture properly is non-trivial. The cost of doing it badly is higher in expectation, factoring in regulatory exposure and litigation risk. The carriers and distributors that are getting this right are the ones treating AI

deployment as a governance project from inception rather than as a technology project that acquires a governance overlay later.

PART IV

# The regulatory perimeter

*The EU AI Act, NAIC Model Bulletin, and UK Consumer Duty as they apply to distribution AI.*

# 15. The EU AI Act and what it means for distribution

The EU AI Act (Regulation 2024/1689) is the first comprehensive AI regulation globally. For insurance distributors operating in or selling into the European Union, it is the dominant compliance reference, and it carries fines up to €35 million or 7% of global annual turnover for the most serious violations — substantially higher than GDPR's ceiling.

## The high-risk classification

The Act classifies AI systems into four tiers: unacceptable risk (prohibited), high-risk (regulated extensively), limited risk (transparency obligations), and minimal risk (largely unregulated). Insurance is explicitly affected at the high-risk tier. Specifically, Annex III lists "AI systems intended to be used for risk assessment and pricing in relation to natural persons in the case of life and health insurance" as high-risk.[4]

Property and casualty AI is in a more ambiguous position. Pricing models for motor, home, commercial property, and casualty lines are not explicitly listed in Annex III, but the Act's broader provisions on automated decision-making affecting essential services may apply. Recent commentary suggests the Commission and EIOPA may issue further guidance, and the prudent posture for distributors is to treat substantial P&C pricing AI as if it were high-risk for governance purposes.

## The May 2026 Digital Omnibus amendments

On 7 May 2026, EU lawmakers reached political agreement on the Digital Omnibus on AI, a package of amendments designed to ease compliance burdens. The key changes for insurance distributors:[4]

- Stand-alone high-risk AI systems (Annex III, including life and health insurance pricing): compliance deadline moves from 2 August 2026 to 2 December 2027 — a 16-month extension.
- Product-embedded high-risk AI systems (Annex I, narrower scope): deadline moves to 2 August 2028.

- AI-generated content watermarking and provenance-labelling: deadline moves to 2 December 2026 — earlier than some original interpretations.
- National AI regulatory sandboxes: deadline for member states to establish them moves to 2 August 2027.

The amendments are provisional pending formal adoption by the Council and Parliament. The Commission has signalled intent to adopt before 2 August 2026 (the current high-risk deadline) to provide certainty. Distributors should plan against the new deadlines but be prepared to default to the original if formal adoption is delayed.[44]

## Obligations on providers and deployers

The Act distinguishes between AI "providers" (entities that develop AI systems and place them on the market) and "deployers" (entities that use AI systems in their professional capacity). Insurance distributors are typically deployers; vendors are typically providers; but the lines blur in specific scenarios — particularly where a distributor substantially modifies a vendor-supplied system, in which case the distributor may be reclassified as a provider with the corresponding broader obligation set.

Key deployer obligations for high-risk AI in insurance:

- Conformity assessment and CE marking (provider obligation, but deployer must verify).
- Technical documentation and instructions for use.
- Risk management system covering the AI system's lifecycle.
- Data governance ensuring quality and bias-mitigation in training and operational data.
- Logging and traceability of system operation.
- Transparency to users, including consumer notification where applicable.
- Human oversight architecture allowing meaningful human intervention.
- Accuracy, robustness, and cybersecurity standards.
- Post-market monitoring with incident reporting to national authorities.

- AI literacy: relevant staff must understand the system's capabilities, limitations, and risks (this obligation applies now, not at the high-risk effective date).

## Solvency II integration

EIOPA has signalled that AI governance is expected to integrate with existing Solvency II Pillar 2 risk-management frameworks and the ORSA process, rather than be implemented as a parallel governance structure. This is an unusually helpful regulatory pragmatism: insurers and MGAs that already have mature Solvency II risk frameworks can extend them to cover AI rather than rebuild from scratch. Distributors with weaker existing risk-management frameworks face a more significant build.

**The practical sequencing for distributors**

Inventory all AI systems in use, including vendor-supplied; classify each under the Annex III risk taxonomy; designate accountable owners; document governance, testing, and human-oversight architecture; integrate with existing Solvency II Pillar 2 frameworks where applicable; plan technical documentation for systems likely to fall into high-risk classification; train relevant staff in AI literacy ahead of the formal effective dates. The carriers and distributors that have completed this work by mid-2027 will be well-positioned for the December 2027 deadline; those starting work in 2027 will be late.

# 16. The NAIC Model Bulletin and the divergence between states

In the United States, insurance regulation is constitutionally a state matter, with the National Association of Insurance Commissioners (NAIC) providing coordination through model laws and bulletins. The NAIC adopted the Model Bulletin on the Use of Artificial Intelligence Systems by Insurers on 4 December 2023, and state adoption has been rapid.

## Adoption status

As of mid-2025, more than 24 US jurisdictions had adopted the NAIC Model Bulletin with little or no material customisation, including Alaska, Connecticut, Delaware, Hawaii, Illinois, Kentucky, Maryland, Massachusetts, Michigan, Nebraska, Nevada, New Hampshire, New Jersey, North Carolina, Oklahoma, Pennsylvania, Rhode Island, Vermont, Virginia, Washington, West Virginia, and Wisconsin.[5][45] Other states have adopted modified versions or related guidance. The total universe of states where carriers should be prepared to demonstrate AI governance practices is now substantially larger than the bulletin-adoption list alone.

## The three pillars

The Model Bulletin is principle-based rather than prescriptive. It does not mandate specific testing methodologies or governance structures in granular detail. What it does is establish that insurers should implement and maintain a written AIS (AI Systems) Program covering three pillars:

- Governance: a documented accountability structure involving business units, actuarial, data science, underwriting, claims, legal, and compliance functions, with senior management or board accountability for oversight.
- Risk management and internal controls: validation, testing, and retesting processes to assess AI outputs, evaluate data quality, and identify potential bias.
- Third-party vendor oversight: due diligence, audit rights in contracts, ongoing monitoring. Responsibility for vendor AI stays with the insurer.

## The Evaluation Tool pilot

Beginning January 2026, 12 participating states are running the NAIC-designed AI Systems Evaluation Tool with four exhibits:[46]

- Exhibit A — Breadth of AI adoption across the enterprise.
- Exhibit B — Governance framework and AIS Program structure.
- Exhibit C — Deep dive on high-risk systems, with current emphasis on agent-facing AI in claims handling, billing disputes, and total-loss decisions.
- Exhibit D — Data source review including proxy-discrimination screening for data used in rating, social-media signals, and aerial imagery.

The pilot is operationalising regulatory examiner capability that has been developing since the bulletin's 2023 adoption. Carriers and MGAs writing business in pilot states should expect more substantive examination of AI governance in 2026 onward; carriers writing in non-pilot states should expect the examination expectations to follow within 12 to 24 months.

## Colorado: the outlier

Colorado has been the most aggressive US state on insurance AI regulation. Three frameworks now apply concurrently to insurers operating in the state: the NAIC Model Bulletin; SB 21-169 (Regulation 10-1-1) requiring quantitative bias-testing; and SB 24-205 (the Colorado AI Act) covering broader algorithmic decision-making, effective from June 2026 with enforcement by the Attorney General.[42][47] Effective 15 October 2025, the SB 21-169 bias-testing scope expanded to include private passenger automobile insurance and health benefit plans alongside the original life-insurance scope.[42]

Colorado's quantitative bias-testing requirement is materially stricter than the NAIC Model Bulletin's principle-based approach. Insurers writing in Colorado must conduct disparate-impact testing on their AI-driven pricing models, even where the insurer does not believe its system discriminates. Annual attestations are required. The second annual attestation cycle (December 2025) has now produced two years of data for the Colorado Division of Insurance, which is signalling that filings that look suspiciously clean will be examined more aggressively going forward.

## New York and other jurisdictions

New York's Department of Financial Services issued a circular letter in 2019 focused on the use of external consumer data in life insurance underwriting, and a 2024 circular letter focused on underwriting and rating fairness more broadly.[48] California, Texas, and several other states have adopted variants of the broader AI-governance framework, though with less prescriptive content than Colorado.

| Jurisdiction | Framework | Status (May 2026) | Key obligation |
|---|---|---|---|
| EU | EU AI Act | Live; Digital Omnibus extending key deadlines to Dec 2027 / Aug 2028 | Conformity assessment for high-risk AI in life/health pricing |
| UK | FCA principles + Consumer Duty + SM&CR | Live; no sector-specific AI rules planned | Fair value, accountability under SM&CR |
| US (Colorado) | SB 21-169 / Reg 10-1-1; SB 24-205 | Live; scope expanded Oct 2025 to auto and health | Quantitative bias-testing; annual attestation |
| US (24+ states) | NAIC Model Bulletin | Live; Evaluation Tool pilot launched Jan 2026 in 12 states | Written AIS Program covering governance, risk, vendor oversight |
| US (New York) | DFS Circular Letters 2019, 2024 | Live | External data use; underwriting and rating fairness |
| US (federal) | No comprehensive federal framework | Pending; state-led patchwork | Sector-specific guidance via FTC, CFPB on specific products |

# 17. The UK approach: principles, Consumer Duty, and SM&CR

The UK has taken a deliberately different approach to AI regulation in financial services from the EU. Where the EU has adopted prescriptive risk-tier requirements, the UK regulators have explicitly chosen principles-based, technology-agnostic, outcomes-focused supervision under existing frameworks.

## The FCA position

In its April 2025 AI Update and subsequent communications, the FCA has been consistent: the regulator does not plan to introduce sector-specific AI rules, on the grounds that existing frameworks — Consumer Duty, SM&CR, SYSC governance rules, outsourcing and operational resilience requirements, and conduct rules — already cover the substantive concerns that AI raises in financial services.[43]

The four primary lenses through which the FCA assesses AI in distribution:

- Consumer Duty: AI-driven products and services must deliver fair value, communicate effectively with target consumers, provide appropriate support, and produce fair outcomes.
- Senior Managers and Certification Regime (SM&CR): accountability for AI-driven decisions cannot be outsourced to a model or vendor. Senior managers remain personally accountable for outcomes.
- SYSC governance: AI systems must have appropriate governance, including documented controls, validation processes, and oversight.
- Operational resilience and third-party rules: AI systems and the third-party vendors that supply them are subject to the same operational-resilience expectations as any other critical business service.

## The Consumer Duty as the dominant lens

PwC's analysis of customer-facing AI in financial services characterises the Consumer Duty as the FCA's primary lens for assessing AI deployments.[33] The duty applies across the AI

lifecycle — from product design through outcomes monitoring — and requires firms to evidence fair-value delivery, foreseeable-harm mitigation, and appropriate consumer understanding of AI-driven processes.

In April 2025 the FCA retired more than 90 "Dear CEO" and portfolio letters, simplifying supervisory communications.[49] Looking forward, the FCA has committed to consult on amendments to the Consumer Duty in the first half of 2026 on its application across distribution chains. For multi-party distribution arrangements — particularly MGA-and-coverholder structures — this is a development worth tracking closely.

## The PRA position

The Prudential Regulation Authority's interest in AI is narrower than the FCA's but consequential. The PRA's focus is on safety and soundness, operational resilience, and the implications of AI dependency on third-party providers (particularly hyperscaler cloud providers and foundation-model developers) for systemic risk. The PRA's recent Financial Stability in Focus publications have begun to map the systemic-risk implications of AI in financial services, including the concentration risk that comes from large numbers of insurers running models on a small number of foundation models from a small number of providers.

## AI Live Testing programme

The FCA's AI Live Testing programme, with a second cohort opening (announced 3 December 2025), provides a structured route for firms to develop AI applications with regulatory engagement. The programme has been particularly used for customer-facing applications: onboarding, customer service chatbots, AI-assisted agents, personalised communications, complaints handling, and affordability assessments.[33]

For insurance distributors, the AI Live Testing route is one of the most cost-effective ways to derisk AI deployment ahead of broader rollout. The regulator's published expectations on participating firms are themselves a useful indication of what the supervisory bar looks like in practice.

## The UK-EU divergence and what it means for distributors

UK distributors selling into the EU face both regimes simultaneously. The compliance architecture must satisfy the EU AI Act's prescriptive requirements where applicable, and the UK regulator's outcomes-focused expectations on the UK book. In practice, this typically means that the EU AI Act requirements are the binding constraint on system architecture, while the UK Consumer Duty is the binding constraint on outcomes monitoring and consumer-facing communication. A well-designed AI governance programme can satisfy both, but it must be designed for both from the start; retrofitting either onto a system built only for the other is expensive.

The Bank of England and FCA's third joint AI and Machine Learning Survey, published November 2024, established important baselines for what regulators expect to see in firm AI governance: 84% of firms have an accountable person for AI; 82% have AI guidelines or best-practice policies; 79% have a data-governance framework; 81% use at least one explainability method, with feature importance (72%) and Shapley values (64%) the most common.[1] Firms operating below these baselines are increasingly visible as outliers.

PART V

# Implications for distribution platforms

*What the operating evidence means for carriers, MGAs, brokers, and the platforms that serve them.*

# 18. Build versus buy, and the integration tax

Every carrier, MGA, and broker evaluating AI in distribution arrives at the build-versus-buy question. The answer is, in nearly every case, both — but the mix matters, and the integration tax is the variable that most often surprises buyers in the first 12 months of deployment.

## The realistic build-versus-buy distribution

Across the credible operator evidence, the dominant pattern of AI procurement in insurance distribution is:

- Buy: foundation models (Anthropic, OpenAI, Google, Microsoft) consumed via hyperscaler APIs. Building proprietary foundation models is uneconomic for any insurer or distributor smaller than the very largest, and arguably for those too.
- Buy: specialist AI-insurance vendors for well-characterised workflows where the vendor has a defensible reference book and a credible data moat. Submission intake (Cytora, Sixfold, Send), pricing (hyperexponential, Earnix, Akur8), fraud (Shift, FRISS), claims (Shift, Tractable). Building these capabilities in-house typically produces a worse result at higher cost.
- Build: integration, orchestration, and data infrastructure. This is the layer where the proprietary advantage is created and where buying off-the-shelf systems is least viable.
- Build: governance, audit trail, and bias-testing infrastructure. Vendor-supplied governance documentation is supporting evidence, not a substitute for the insurer's own programme.
- Build or buy (case by case): customer-facing conversational AI. The trade-off between brand-consistency (favouring build) and time-to-market (favouring buy) is genuinely context-dependent.

## The integration tax

The cost of integrating AI vendors with existing core systems — policy administration, claims administration, billing, agency management — is the most commonly

underestimated line in the AI business case. A representative pattern: a £500k vendor licence produces £1.2m of first-year integration cost, £400k of annual ongoing integration maintenance, and £200k of bias-testing and audit-trail cost. The economics still work for most well-chosen AI deployments, but the integration tax determines the break-even point.

Three integration-tax drivers are worth flagging explicitly:

- Core system interface immaturity. The major core platforms (Guidewire, Duck Creek, Sapiens, Insurity) have API coverage that varies in completeness across modules. Integrations to mature modules are tractable; integrations to less-mature modules can require significant custom development.
- Data normalisation cost. AI vendors require structured, normalised data inputs. Most core systems do not produce those natively. The data-normalisation work between core and AI is often the dominant integration cost line, and it is the work that should not be skipped, because data quality dominates AI model performance.
- Governance integration. Adding AI to an existing governance, change-management, and audit-trail framework requires both technical and process work. The technical work is usually completed first; the process work is what determines whether the governance is robust in production.

## The platform argument

Where the integration tax for individual point-vendor AI deployments is high, there is a corresponding argument for distribution platforms that absorb the integration cost by orchestrating multiple AI capabilities through a single architectural layer. This is structurally the most attractive position in the insurance technology stack from a platform-economics standpoint — the platform captures the integration value, while AI vendors compete to be best-in-class within the platform's orchestration architecture.

The McKinsey investor analysis frames AI infrastructure providers — "vendors enabling connectivity across data layers, models and automated agents" — as a structurally favoured category for capital allocation, distinct from AI feature suppliers.[7] The same framing applies to distribution platforms that play this role within the carrier-to-customer flow rather than within the AI tooling layer specifically.

**The discipline that distinguishes buyers who succeed**

Buyers who succeed treat AI procurement as a five-year capability decision, not a 12-month software purchase. They specify the integration architecture before selecting the vendor. They budget for the full integration tax — typically 2–4x the licence cost in year one. They build governance into procurement contracts (audit rights, model documentation, bias-testing cooperation). They sequence deployments to start with low-regulatory-risk workflows (intake, internal triage) before moving to higher-risk ones (pricing, denial decisions). And they treat the data-flywheel question — covered in the next section — as a first-class procurement criterion rather than an afterthought.

# 19. The data flywheel: the underrated moat

If there is a single concept underexplained in the trade-press AI coverage and overrepresented in the operating reality of successful AI in distribution, it is the data flywheel. The flywheel is the feedback loop between AI deployment, data accumulation, model improvement, and operational advantage. It is the structural moat that distinguishes durable AI advantage from temporary feature parity.

## How the flywheel works in distribution

The flywheel mechanism in insurance distribution typically operates as follows:

- AI is deployed against a specific workflow (intake, pricing, claims, retention).
- The deployment generates structured data about customer behaviour, risk characteristics, and outcomes that previously existed only as unstructured information or did not exist at all.
- That structured data feeds back into model improvement, segmentation refinement, and pricing precision.
- Improved models produce better customer outcomes (more competitive prices, faster decisions, more accurate fraud detection), which attracts more business.
- More business produces more data, which closes the loop.

Lemonade's published positioning is explicit: "Faster growth expands our data advantage, which sharpens our AI-powered segmentation and pricing models."[23] The structural argument is sound, and Lemonade's own data — the pet insurance line growing 55% to $439 million as the data advantage compounds — is supportive evidence.

## Where the flywheel is most powerful in distribution

Three structural conditions are required for the flywheel to operate effectively:

- The AI deployment must touch the moment of customer interaction directly, not just the back-office work that surrounds it. Submission triage produces structured data about broker behaviour; quote intake produces structured data about

customer intent. Pure back-office automation (claims-paperwork extraction, regulatory reporting) does not generate the same flywheel.

- The data captured must be operationally consequential for the model. Logging the data is not enough; the model must be retraining or refining on it. Static models that never update do not benefit from the flywheel.
- The volume must reach a threshold of statistical significance for the model improvements to be measurable. Below that threshold, the apparent improvements are within noise.

## The MGA opportunity

McKinsey's analysis identifies MGAs as the segment of insurance distribution best positioned to capture flywheel advantage, on the grounds that MGAs combine three things rarely co-located elsewhere: direct distribution relationships, proprietary specialty data, and the agility to deploy AI quickly.[50] The McKinsey thesis: "MGAs that combine strong relationships with data ownership and advanced AI use will differentiate themselves most sharply."

This thesis is consistent with the observed growth in MGA premium volumes (US direct premium nearly doubled from $47bn in 2020 to $97bn in 2024, an approximately 14% annual growth rate, well above the broader insurance market). The structural advantage MGAs hold in the flywheel is real; the operational requirement to convert that advantage into durable competitive position is significant; the procurement decisions made between 2025 and 2027 will determine which MGAs participate in that advantage and which do not.

## Implications for distribution platforms

The flywheel question reframes the platform-versus-vendor distinction. A distribution platform that orchestrates AI capabilities but does not capture the data flywheel itself is a thin-margin intermediary in a commodity market. A distribution platform that orchestrates AI capabilities and captures the data flywheel — by sitting at the moment of customer or broker interaction, owning the structured data that interaction produces, and feeding that data back into model improvement — is in a structurally different competitive position. The

platform layer is where the flywheel can be made durable across multiple AI capabilities, even as the individual AI vendors below the platform layer rotate over time.

# 20. A short note on the cXc view

cXc — a trading name of conXchange Pvt Ltd — operates at the distribution layer that this paper has spent most of its pages examining. The brief note below sets out where cXc sits on the questions the paper has covered, for readers who want to understand our perspective. The body of the paper is independent of this section; the conclusions stand whether or not cXc's view matters to a particular reader.

## Our operating thesis

Insurance distribution is the part of the value chain where the gap between current technology capability and current operational practice is largest. The carriers, MGAs, brokers, and reinsurers who close that gap first will hold structural advantage for the rest of the decade. The technology to close the gap exists. The integration architecture, governance infrastructure, and data-flywheel design to operationalise it do not yet exist at most operators in the market.

cXc is built on the premise that this layer — the integration, orchestration, and governance layer that sits between the AI vendors and the regulated insurance operators — is where the most useful work in insurance technology can be done in the 2026 to 2030 window. We do not build foundation models. We do not compete with the specialist AI-insurance vendors named throughout this paper. We integrate them, orchestrate them, and supply the governance and data-flywheel infrastructure that lets carriers and distributors operate them safely and durably.

## The specifics we care about

Three operating practices that we believe distinguish durable AI deployments in distribution from temporary ones:

- Governance integrated from procurement, not retrofitted post-deployment. The carriers and MGAs who do this will be ready for the EU AI Act December 2027 deadline, the NAIC Evaluation Tool examinations, and the Consumer Duty supervisory reviews. The ones who do not will be replatforming under pressure.

- Data flywheel design at the platform layer, not at the AI vendor layer. The data must accumulate in a place that is durable across AI vendor rotation, because no specific AI vendor will be the right answer for five years; the data infrastructure must outlive the specific vendor choices made on top of it.
- Human-in-the-loop architecture as a serious operating system, not a documentation overlay. The architectures that survive regulator scrutiny are the ones where the human reviewer is triggered by informative signals, has real authority to override, and produces a full audit trail of decisions. This is more expensive to build than to claim, and the difference is visible in market-conduct examinations.

## Where to engage

Readers who want to discuss any of the territory this paper has covered — the use-case maturity assessments, the regulatory implications, the data-flywheel architecture, or specific operational questions about AI deployment in their own distribution — can reach cXc at the contact details on cxc.services. The research team responds personally.

**On the use of AI in producing this paper**

This paper was researched and drafted using a combination of human research and large-language-model assistance. Every statistic cited has a named primary source in the references section. Every operator named is identified with a publicly verifiable example or published statement. Where the evidence is mixed or where vendor claims could not be independently verified, the paper says so. cXc Research is the human-accountable author of the work.

# References

All sources accessed May 2026. Inline citation numbers map to entries below.


**[1]** Bank of England and Financial Conduct Authority, "Artificial intelligence in UK financial services — 2024" (November 2024). https://www.bankofengland.co.uk/report/2024/artificial-intelligence-in-uk-financial-services-2024

**[2]** Conning, 2025 AI and Insurance Technology survey, cited in MoneyGeek, "How AI Is Changing Insurance in 2026" (March 2026). https://www.moneygeek.com/insurance/how-ai-is-changing-insurance/

**[3]** Lloyd's Market Association, "Over one-third of London market firms now actively using AI" (24 April 2025). https://lmalloyds.com/over-one-third-of-london-market-firms-now-actively-using-ai/

**[4]** Latham & Watkins, "AI Act Update: EU Resolves to Change Rules and Extend Deadlines" (May 2026). https://www.lw.com/en/insights/ai-act-update-eu-resolves-to-change-rules-and-extend-deadlines

**[5]** Quarles & Brady, "Nearly Half of States Have Now Adopted NAIC Model Bulletin on Insurers' Use of AI" (March 2025). https://www.quarles.com/newsroom/publications/nearly-half-of-states-have-now-adopted-naic-model-bulletin-on-insurers-use-of-ai

**[6]** Financial Conduct Authority, "AI and the FCA: our approach" (February 2026). https://www.fca.org.uk/firms/innovation/ai-approach

**[7]** McKinsey & Company, "AI in insurance: Understanding the implications for investors" (February 2026). https://www.mckinsey.com/industries/financial-services/our-insights/ai-in-insurance-understanding-the-implications-for-investors

**[8]** Insurance Business, "AI is accelerating in insurance — are you ready?" (April 2026), citing Grant Thornton 2025 executive survey. https://www.insurancebusinessmag.com/us/news/technology/ai-is-accelerating-in-insurance--are-you-ready-573111.aspx

**[9]** TheStreet, "McKinsey says AI could reshape how you buy insurance" (March 2026), citing Bank of America commission analysis. https://www.thestreet.com/personal-finance/mckinsey-says-ai-could-reshape-how-you-buy-insurance

**[10]** Intelliarts, "Case Study: Predictive Lead Scoring Software for Insurance". https://intelliarts.com/success-stories/predictive-lead-scoring/

**[11]** Smartlead, "Case Studies: Companies That Improved Conversions with AI Lead Scoring" (March 2026), referencing Progressive Insurance ML app feature. https://www.smartlead.ai/blog/case-studies-companies-that-improved-conversions-with-ai-lead-scoring

**[12]** Lido, "Best Underwriting Automation Software in 2026". https://www.lido.app/blog/best-underwriting-automation-software

**[13]** SortSpoke, "ACORD Forms Processing: Extract Data 5X Faster". https://sortspoke.com/solutions/document-types/acord-forms

**[14]** Insurance Journal, "Digital Lloyd's Syndicate Ki Becomes Separate Company Within Fairfax Group" (January 2025). https://www.insurancejournal.com/news/international/2025/01/24/809511.htm

**[15]** Insurance Journal, "Digital Lloyd's Syndicate Ki Expands Follow Capacity" (April 2026). https://www.insurancejournal.com/news/international/2026/04/28/867546.htm

**[16]** Lemonade Inc., Form 10-K Annual Report for fiscal year 2025, filed February 2026. https://www.stocktitan.net/sec-filings/LMND/10-k-lemonade-inc-files-annual-report-33aac5b74a32.html

**[17]** Galileo AI, "Best LLMs for AI Agents in Insurance" (August 2025), referencing Klarna AI replacement reversal. https://galileo.ai/blog/best-llms-for-ai-agents-in-insurance

**[18]**"Agentic AI for Commercial Insurance Underwriting with Adversarial Self-Critique", arXiv preprint 2602.13213. https://arxiv.org/pdf/2602.13213

**[19]**"Benchmarking Agents in Insurance Underwriting Environments", arXiv preprint 2602.00456. https://arxiv.org/pdf/2602.00456

**[20]**Digital Insurance, "Hyperexponential launches AI-based underwriting platform" (September 2025). https://www.dig-in.com/news/hyperexponential-launches-ai-based-underwriting-platform

**[21]** Underwriting Software, "Insurance Underwriting Software — Automate Submission Processing". https://www.underwritingsoftware.co/

**[22]**Cytora, customer case studies (Zurich, Markel). https://cytora.com/

**[23]** Perspective AI, "How Conversational AI Made Lemonade the Fastest-Growing AI Insurance Company in Pet Insurance" (February 2026), citing Lemonade Q4 2025 shareholder letter. https://getperspective.ai/blog/lemonade-case-study-conversational-ai-insurance

**[24]** Mordor Intelligence, "Embedded Insurance Market Outlook: 30%+ CAGR Forecast Through 2031" (March 2026). https://www.prnewswire.com/news-releases/embedded-insurance-market-outlook-30-cagr-forecast-through-2031-online-and-api-first-placements-held-76-38-share-in-2025--reports-mordor-intelligence-302712506.html

**[25]**Coinlaw, "Embedded Insurance Industry Statistics 2025". https://coinlaw.io/embedded-insurance-industry-statistics/

**[26]**Mordor Intelligence, "Europe Embedded Insurance Market Size & Growth to 2031" (January 2026), referencing Zurich-Ominimo April 2025 partnership. https://www.mordorintelligence.com/industry-reports/europe-embedded-insurance-market

**[27]**EZLynx, "Insurance Data Analytics for Operational Efficiency". https://www.ezlynx.com/blog/posts/insurance-data-analytics-operational-efficiency/

**[28]** Salesforce, "Best Insurance Agency Management & Brokerage Software". https://www.salesforce.com/financial-services/insurance-agency-management-software/

**[29]**Insurity, "Insurity Predict: Predictive Analytics in Insurance". https://insurity.com/insurity-analytics/predictive-analytics

**[30]**Intelliarts, "Customer Churn Analysis in Insurance with Machine Learning" (January 2026). https://intelliarts.com/blog/machine-learning-for-customer-churn-prediction-in-insurance/

**[31]** Schalck, "Innovative Techniques to Predict Churn in the French Insurance Industry: Integration of Machine Learning With the Grabit Model", Journal of Forecasting (Wiley, October 2025). https://onlinelibrary.wiley.com/doi/10.1002/for.70057

**[32]** "Applying economic measures to lapse risk management with machine learning approaches", arXiv preprint 1906.05087. https://arxiv.org/pdf/1906.05087

**[33]** PwC UK, "Scaling customer-facing AI: unlocking better outcomes and Consumer Duty compliance". https://www.pwc.co.uk/industries/financial-services/understanding-regulatory-developments/scaling-customer-facing-ai-unlocking-better-outcomes-and-consumer-duty-compliance.html

**[34]** Life Insurance International, "Lloyd's syndicate Ki becomes stand-alone entity within Fairfax Group" (January 2025). https://www.lifeinsuranceinternational.com/news/ki-standalone-entity-fairfax/

**[35]** Ki Insurance, "Future of AI in underwriting" (September 2024). https://ki-insurance.com/news/future-of-ai-in-underwriting

**[36]** Shift Technology, customer case studies. https://www.shift-technology.com/resources/case-studies/author/shift-technology

**[37]** Shift Technology, "Shift Technology Launches Shift Claims to Power Claims Transformation with Agentic AI" (September 2025). https://www.shift-technology.com/resources/news/shift-technology-launches-shift-claims-to-power-claims-transformation-with-agentic-ai

**[38]** Insurance Innovation Reporter, "Shift Technology Launches AI Claims Platform" (September 2025). https://iireporter.com/shift-technology-launches-ai-claims-platform/

**[39]** Health Affairs, "The AI Arms Race In Health Insurance Utilization Review" (2025). https://www.healthaffairs.org/doi/10.1377/hlthaff.2025.00897

**[40]** Browne Jacobson, "AI hallucinations" (May 2025), referencing Armilla AI insurance launch and Air Canada chatbot case. https://www.brownejacobson.com/insights/the-word-may-2025/ai-hallucinations

**[41]** Raizner Slania, "Lemonade Insurance's AI Technology Could Lead to Wrongful Denial of Claims" (May 2025). https://www.raiznerlaw.com/insights/lemonade-insurances-ai-technology-could-lead-to-wrongful-denial-of-claims/

**[42]** WaterStreet Company, "Colorado SB 205 Rules on Insurance AI" (April 2026). https://www.waterstreetcompany.com/colorado-sb-205-insurance-ai/

**[43]** A-Team Insight, "FCA AI Update 2025: How the Regulator is Embedding AI Oversight into UK Financial Rules" (September 2025). https://a-teaminsight.com/blog/fca-ai-update-2025-how-the-regulator-is-embedding-ai-oversight-into-uk-financial-rules/

**[44]** Travers Smith, "EU agrees to delay key AI Act compliance deadlines" (May 2026). https://www.traverssmith.com/knowledge/knowledge-container/eu-agrees-to-delay-key-ai-act-compliance-deadlines/

**[45]** WaterStreet Company, "What the NAIC Model Bulletin Means for Insurance AI" (April 2026). https://www.waterstreetcompany.com/what-the-naic-model-bulletin-means-for-insurance-ai/

[46] VerifyWise, "NAIC Model Bulletin on Use of AI Systems by Insurers". https://verifywise.ai/ai-governance-library/sector-specific-governance/naic-ai-model-bulletin

[47] VerifyWise, "Colorado SB 21-169 compliance playbook for insurers" (April 2026). https://verifywise.ai/es/blog/colorado-sb-21-169-compliance-playbook-for-insurers

[48] McDermott Will & Emery, "State Regulators Address Insurers' Use of AI: 11 States Adopt NAIC Model Bulletin" (September 2025). https://www.mwe.com/insights/state-regulators-address-insurers-use-of-ai-11-states-adopt-naic-model-bulletin/

[49] Skadden, "FCA Clarifies Aspects of the Consumer Duty" (October 2025). https://www.skadden.com/insights/publications/2025/10/fca-clarifies-aspects-of-the-consumer-duty

[50] McKinsey & Company, "AI in insurance: Understanding the implications for investors" (February 2026), section on MGAs. https://www.mckinsey.com/industries/financial-services/our-insights/ai-in-insurance-understanding-the-implications-for-investors

**Note on methodology**

Where multiple credible sources reported the same statistic, this paper has prioritised primary sources (regulator publications, company financial filings, peer-reviewed research). Where only secondary sources were available, the paper has cited the closest credible secondary source to the primary data. Where a statistic was reported by a vendor and could not be independently verified, the paper has identified the source as the vendor rather than presenting the statistic as third-party-verified. Readers who want to verify specific numbers should consult the linked primary sources.

cXc

conXchange Pvt Ltd · cxc.services