# AI agents inside the insurer

## What they can read, what they can reach, and who decides

**IN BRIEF**

01 Insurers have put AI in front of customers and staff at scale, and the first agents that act are in production: taking in claims, checking cover, preparing submissions. They wait for a person where money leaves the company and where a customer is refused.

02 Insurance work is documents first and systems second. An agent has to read the claim pack and the policy wording, and then reach the policy and claims systems. Most tools do one or the other.

03 Public catalogues of ready-made connections do not contain an insurer's systems of record. Those are reached through the APIs the insurer already runs.

04 An insurer needs one place where every tool an agent can call is listed, every call is checked against policy, a payment above the insurer's own threshold waits for a named person, and the record can be put in front of a supervisor.

05 All of this can run on infrastructure the insurer controls. The agents the core-system vendors now sell run in the vendors' clouds.

IN THIS PAPER

01 Where insurers are today 2
02 The systems of record are in nobody's catalogue 3
03 Three doors, one gate 4
04 Six ways into an insurance system 5
05 What an agent may do 6
06 Seven questions before an agent touches a claim or a policy 6
07 What regulators already ask 7
08 Where to start 8

PART 01

# Where insurers are today

**82%**
of Ping An's customer service volume handled by AI

**25 minutes**
for an AIA health claim in South Korea, from over two days

**7 agents**
on one kind of claim at Allianz in Australia

**60 million**
documents a year at Daman in Abu Dhabi

**The scale is in service and in simple claims.** Ping An says AI now handles 82% of its customer service volume, that its express service resolves 91% of requests without a person, and that 95% of its life policies are underwritten within seconds. AIA says a health claim in South Korea that took more than two days is now completed in under 25 minutes, on a system it built itself because, it says, no vendor could meet its requirements.

**The documents come first.** In Abu Dhabi, the health insurer Daman handles more than 60 million documents a year, and its group's new system is expected to turn most of them into structured data and to flag what is missing when a file is submitted. Before any agent decides anything, something has to read the pack.

**Agents that act are narrower, and already in production.** Allianz put seven agents to work in Australia on one kind of claim: food spoiled in a power cut after a storm, under 500 Australian dollars. Allianz names them: planner, cyber, coverage, weather, fraud, payout and audit. Suncorp, also in Australia, said in June 2026 that five agents were going into production in its claims process: two that take the first report of a loss, at least one of them by voice, and others that check cover, dispatch a repairer and calculate a settlement. IAG counts 92 generative AI uses in production.

**The line is drawn where money leaves and where a customer is refused.** Allianz: "By design, payout decisions are never automated." Suncorp says every claim an agent would reject is to go through a mandatory human review, and that a person takes over when a report gets too complex. AIA does not reject a flagged claim automatically; it alerts a human assessor. In Hong Kong, Prudential's underwriting assistant gives a consultant a preliminary outcome in minutes, before any application is filed. In the UAE, Sukoon has announced an agent that registers a complete claim by itself and drafts every email for a member of staff to check before it is sent.

**Below that line, the simple goes straight through.** Ping An's express service resolves nine requests in ten with no person involved, and Sukoon's agent is to register a complete claim by itself.

So the question for an insurer is not whether an agent may ever act alone. It is who sets the threshold, where it is written down, and whether it holds.

PART 02

# The systems of record are in nobody's catalogue

**327**
servers in Docker's public catalogue on 19 November 2026

**0**
of them an insurance system

**5 in 100**
insurance products whose API description anyone can download

**1**
policy administration vendor publishing its own MCP server

The Model Context Protocol, MCP, has become the common way to give an agent a tool. A system is wrapped once as an MCP server, and any agent can use it.

The largest public catalogue of these servers is Docker's. We read every entry. On 19 November 2026 it held 327 servers. They cover databases, document stores, ticketing, monitoring, maps, weather and web search.

**Not one is an insurance system.** There is no policy administration system, no claims system, no billing system, no rating engine, no agency management system, no reinsurance system and no health claims exchange. The word "insurance" appears in none of the 327 descriptions.

The insurance vendors' own efforts do not fill the gap yet. Among the vendors of policy administration systems we found one that publishes its own server and at least two that have announced one. Each runs in the vendor's cloud.

The same vendors have gone a step further than the banking vendors. In the eighteen months to November 2026 nearly every one of them launched agents of its own: for the first report of a loss, for claim summaries, for underwriting submissions, for brokers. They are sold as part of the vendor's cloud service. An insurer that runs its core on its own infrastructure, or that keeps health data in the country by law, cannot use them as they are.

There is a second fact underneath this one. **Most insurance systems describe their APIs for customers only.** We found five products out of about a hundred whose API description anyone can download. For the rest, and even for those five, the file an insurer needs is the one generated from its own installation, with its own lines of business and its own fields. Nobody else can connect an agent to it in advance.

So the route to an insurer's systems of record does not run through a catalogue. It runs through the insurer's own APIs, described by the insurer's own files.

PART 03

# Three doors, one gate

An insurer's work has paper at the front, systems at the centre and services at the edge. An agent platform needs a door to each.

**The first door is for documents.** A claim arrives as a form, invoices, a medical report, photographs. A submission arrives as a broker's email with attachments. On Katonic a user gives the agent the files in the conversation, and the agent reads them. The standing documents, the policy wordings, tables of benefits, underwriting guidelines, claims procedures and treaty terms, are indexed once as knowledge, with each reader's permissions kept, so an agent answers from the wording and cites the clause.

One rule belongs here. **Where the wording is silent, the agent says so.** An answer about cover has four possible states: covered, covered with conditions, not covered, and needs a person. An agent that infers cover from a document that does not grant it has made the insurer a promise.

**The second door is the catalogue, for the edge.** Katonic's catalogue of MCP servers is drawn from Docker's and carries Katonic's own additions. A server is deployed as a container inside the insurer's own cluster. This is how an agent reaches a database of policy and claims records, a document store, a map or weather service for a property risk, or a web search for an investigation.

One caution applies to any catalogue. A container that runs inside the insurer may still call a service outside it. Only a server whose target the insurer hosts keeps everything inside.

**The third door is the insurer's own APIs, for the centre.** Most systems an insurer runs can describe their API in a standard file, an OpenAPI or Swagger description. Katonic imports that file, and each operation in it becomes a tool an agent can call.

- The file is checked before anything is created. The insurer sees how many tools it will produce and any warnings.
- Each tool carries the operation's own inputs, so the model is told what the API expects and does not guess.
- Operations the agent should not have can be switched off one by one. A claims API has hundreds; a triage agent needs ten.
- A tool can be tried with real values, or as a dry run that shows the exact request and sends nothing.
- Each connection has its own sign-in, including the client id and secret that most insurance systems expect from another system, a limit on calls per minute, and a time limit.
- A document can travel with a call: attached to a claim, or fetched from the policy system.

**All three doors lead to one gate.** A document, a tool from a catalogue server and a tool from an imported API sit under the same policies, with the same record.

THREE DOORS, ONE GATE

| Door 01 | Door 02 | Door 03 | All three |
|---|---|---|---|
| **Documents.** | **The catalogue.** | **Your own APIs.** | **One gate.** |

An agent builder sees tools. A risk officer sees one inventory.

PART 04

# Six ways into an insurance system

Not every system has a modern API. The work of giving it one is done in products the insurer already owns, and it follows a small number of patterns.

| | THE SITUATION | WHAT THE INSURER DOES | WHAT THE AGENT PLATFORM DOES |
|---|---|---|---|
| 01 | **The vendor hosts an MCP server for the system** | Registers the agent platform as a client and issues a credential | **Connects to it, and sets a policy on each tool** |
| 02 | **The system produces its own API description** | Exports the file from its own installation and names the operations agents may use | **Imports the file** |
| 03 | **The vendor documents the API for customers only** | Hands over its copy of the file and a credential | **Imports the file** |
| 04 | **The insurer's API gateway already fronts the system** | Registers the agent platform as one more consumer | **Imports the gateway's file, or connects to the gateway's own MCP endpoint where it has one** |
| 05 | **An older policy system: SOAP, a mainframe transaction, a batch file** | Puts a REST face over it in the integration product it already runs | **Imports the file that product generates** |
| 06 | **A system with only a screen, such as a portal** | Records the steps in a screen-automation product | **Connects to that product's tool, for reading only** |

Two things follow. Nothing is replaced: the agent comes in through the same door as every other consumer of the insurer's APIs. And a life insurer's oldest book is usually behind pattern 5.

That work is the insurer's, and it is worth doing once for every consumer, not for agents alone.

PART 05

## What an agent may do

Reaching a system is half of it. The other half is deciding what an agent may do there, and being able to show the decision afterwards.

**Every call is checked before it runs.** On Katonic the order is fixed: is this user allowed at all, do the inputs carry anything that looks like an injection attack, and what does the insurer's policy say about this tool.

**A policy gives one of four answers.** Allow it. Allow it and record that it was watched. Hold it until a person approves. Block it. A policy can cover one tool, every tool on a server, or every tool, and it can apply to the whole company or to one team. Where two policies meet, the stricter one wins.

**A tool that changes something waits.** Each imported operation is rated by what it does: reading is low risk, changing is medium, deleting is high. When an agent is published, any tool that changes something and has no policy of its own is set to wait for approval. An insurer can relax that for a tool: letting a complete claim register itself, say, or a small payment go through. It has to decide to, and the decision is on record.

**If the policy cannot be read, the call is refused.** A fault in the control must not become permission.

**The agent acts as the person using it.** Where an insurer's API checks who is calling, Katonic passes the signed-in user's identity with each call. The value comes from the platform, not from the model, so an agent cannot be talked into acting as someone else. A handler's agent sees the claims that handler may see.

**Every call leaves a record:** who asked, which agent, which tool, what the policy decided and who approved.

**There is always a way to stop.** An agent, a tool or a whole server can be switched off at once.

PART 06

## Seven questions before an agent touches a claim or a policy

These are worth asking of any agent platform, ours included.

| | |
|---|---|
| 01 | **Is there one list of every agent in the company, and of every tool each one can call?** |
| 02 | **Which of those tools change something, and who decided each one may run?** |
| 03 | **Above what amount does a payment need a person, who set that amount, and where is it written?** |
| 04 | **Whose identity does the agent act under when it calls a system?** |
| 05 | **What does the agent do when the policy wording is silent?** |
| 06 | **Where does the model run, and what leaves the insurer's network on each request?** |
| 07 | **Can every call be shown to a supervisor, with the decision that allowed it?** |

PART 07

# What regulators already ask

Few rules are written for AI agents yet. Singapore's central bank, which supervises insurers, issued guidelines on AI risk management on 7 October 2026 that bring agents into scope, and they take effect from 7 October 2027. It has said it will consult in 2027 on further guidance for agents. Elsewhere agents fall under rules an insurer already lives with: technology risk, outsourcing, conduct and data protection.

Across the regulators we read, the same expectations recur. The board stays accountable, including for AI bought from a vendor. People oversee a system in proportion to its risk. The insurer keeps an inventory of its AI systems. A system is tested before production and monitored after, with a way to suspend it. Records are kept. Third parties are assessed, and reliance on a single provider is watched.

**What one regulator already says**

The Central Bank of the UAE, which supervises insurers, issued a guidance note on 11 February 2026 on the use of AI by licensed financial institutions, "including insurance providers". It reads like a specification for section 5 of this paper.

- It names three ways a person can oversee a system: in the loop, where the person approves or rejects each outcome; on the loop, where a person monitors and can intervene; and out of the loop.
- The third is for "low-risk, non-material processes" only, with controls in place.
- It gives an "insurance claim" as its example of a high-impact decision.
- An institution should be able at all times, clearly and immediately, to stop using an AI system.
- It should keep an inventory of all its AI models, and "should not employ AI models that they have no control over".

MARKET BY MARKET

**The Gulf.** The UAE note also asks for testing for bias once a year or whenever a model changes, for audit rights over AI suppliers, and for personal data to be handled in line with the rules on keeping data in the country. In Dubai, the health authority's policy on AI binds health insurers in the Emirate: where an AI system may significantly affect people, a designated person must be able to override it or reverse its decision. Health data in the UAE sits under its own law.

**Malaysia.** Bank Negara supervises insurers and takaful operators under the same technology-risk policy as banks, reissued on 25 September 2026. Its guidance names agentic AI as an emerging technology. An institution that first puts an emerging technology into a critical system consults the Bank beforehand, with a risk assessment and an outside review, whether the system runs in a cloud or in its own data centre. Public cloud is permitted and brings further steps.

**Australia.** The prudential regulator wrote to every regulated entity on 30 April 2026. It expects an inventory of AI tools and uses, and fallback processes where AI supports critical operations, and it observed that identity and access management has not yet caught up with non-human actors such as AI agents.

So running agents on the insurer's own infrastructure does not take away the duty to assess, and in Malaysia to consult. It takes away the questions about where the infrastructure is, who else shares it, and which country's law reaches the data. And the controls in section 5 are the evidence a risk assessment asks for: a list of agents and their tools, a threshold with a named approver, a record of each decision, and a way to stop.

PART 08

# Where to start

Pick work where reading is slow, the decision stays with a person, and the documents are already in hand.

## THREE PLACES THAT FIT

### Policy questions

Staff, brokers or members ask about cover and get answers from the wording itself, with the clause cited and “needs a person” where the wording is silent. Nothing is changed, so there is nothing to approve.

### Claim pack completeness

The agent compares a submitted pack with the checklist for that kind of claim and lists what is missing. A handler reviews the list before anyone is contacted.

### Coverage check

The agent sets a loss against the wording: the clause that grants cover, the exclusions, the limits and the deductible, each quoted. The adjuster decides.

## THEN, IN ORDER

01 Name the documents and the systems the agent needs, and which of the six ways each system is reached.

02 Decide the policy for every tool, and the threshold for every payment, before the agent is built.

03 Build it with the people who do the work today.

04 Run it on invented data, and then beside the manual process.

05 Measure it against that process: time, and how often a person overturned it.

06 Go live for one team and one line of business.

A fuller list of 31 agents across ten areas is published with this paper.

### About Katonic AI

Katonic AI is the enterprise platform for serving models and building, running and governing AI agents. It runs on your machines, with your models, under your administrators. Katonic is listed in the Red Hat Ecosystem Catalog and runs on Red Hat OpenShift. The platform has run in production for about 1,100 staff of a life insurer in Asia, on the insurer’s own cluster, inside its network, and it is in production at a top-tier bank in Asia, air-gapped, inside the bank’s own perimeter. It was founded in Sydney in 2020 by Prem Naraindas, with its global hub in Dubai and teams in Sydney, Mumbai and Dubai.

KATONIC.AI

## SOURCES

All read on 20 November 2026.

01 Insurance Business Asia, report on Ping An’s nine-month results, 29 October 2026.

02 (Re)in Asia, interview on how Asia’s life insurers use generative AI, 16 April 2025 (AIA).

03 Allianz, “When the storm clears, so should the claim queue”, 3 November 2025. allianz.com

04 iTnews, “Suncorp to have AI agents in insurance claims process as soon as this month”, 22 June 2026.

05 iTnews, “IAG drives AI more deeply across its operations”, 13 May 2026.

06 Insurance Business Asia, “Hong Kong insurers take AI closer to the point of sale”, 10 September 2026 (Prudential Hong Kong).

07 Sukoon Insurance, announcement of an agentic AI solution for first notice of loss, as published by Zawya, undated.

08 Abu Dhabi Media Office, announcement of PureHealth’s document system for Daman, 18 September 2026.

09 Docker MCP registry, 327 servers, read in full on 19 November 2026.

10 Central Bank of the UAE, guidance note on the responsible adoption and use of AI and machine learning by licensed financial institutions, issued 11 February 2026. rulebook.centralbank.ae

11 Monetary Authority of Singapore, Guidelines on Artificial Intelligence Risk Management, and media release of 7 October 2026. mas.gov.sg

12 Dubai Health Authority, Policy “Artificial Intelligence in the Healthcare”, DHA/HISHD/PP-09. dha.gov.ae

13 Bank Negara Malaysia, Risk Management in Technology, reissued 25 September 2026, and its frequently asked questions, updated 1 July 2026. bnm.gov.my

14 Australian Prudential Regulation Authority, letter to industry on artificial intelligence, 30 April 2026. apra.gov.au