The Checklist

A MARVELX FIELD GUIDE FOR CLAIMS LEADERS

Six criteria for evaluating *AI claims agents.*

A Head of Claims now sees the same demo from every vendor. A clean approval, a confident number, a screen that never shows the claim that goes wrong. The hard questions arrive later, from an auditor, a regulator, or a policyholder disputing a settlement.

We built this with claims teams who had sat through a dozen demos and still could not say which vendor would hold up in an audit. Six questions did the sorting. A serious vendor clears all six.

01 **Explainability**	04 **Core-system integration**
02 **Human oversight**	05 **Proof over promises**
03 **Data residency and compliance**	06 **Accuracy, defined**

Show the why

CRITERION 01 OF 06

01 Explainability

Every decision reproducible on demand.

When a policyholder disputes an outcome, you have to show the reasoning behind it. A confident approval you cannot explain is a liability. Your handlers are the ones who absorb the risk.

WHAT GOOD LOOKS LIKE

- A trace of the inputs, the rules applied, and the outcome for any single decision.
- Reachable by a handler directly, not through a support ticket.
- Written so an auditor can read it, not only an engineer.

ASK THE VENDOR

Can you reproduce the full reasoning for one decision we pick, live?

REJECT IF

The trace needs a support ticket, or “the model decided” is the answer.

FIG. 01 AUDIT-READY TRACE

01	02	03
Inputs	Rules applied	Outcome

HOW MARVELX APPROACHES THIS

Audit-ready output documents every step of a claim decision, so the trace exists before anyone asks for it.

Keep a person

CRITERION 02 OF 06

02 Human oversight

Automation with a clear route back to a person.

The system should handle what does not need judgment and route the rest to a handler with full context. Oversight is a requirement your team sets, not a limitation of the product. Automation with no way back will make the wrong call on the claims that most need a person.

WHAT GOOD LOOKS LIKE

- An escalation threshold documented and agreed before go-live.
- Escalated claims arrive with the evidence and the proposed decision attached.
- Handlers review in minutes rather than reopening the file.

ASK THE VENDOR

Where is the escalation threshold documented, and what does a handler see?

REJECT IF

Automation has no clear route back to a person.

FIG. 02 HOW THE AGENT HANDLES EVERY CLAIM

A claim arrives

AI CLAIMS AGENT, ON EVERY CLAIM

Reads the file, validates the policy, and chases any missing evidence before a decision is made.

CLEAN, LOW-AMBIGUITY	→	Straight-through, full audit trail
LOW-CONFIDENCE OR EDGE	→	Handler receives the evidence gathered and a drafted decision

HOW MARVELX APPROACHES THIS

The agent works both paths. Clean claims clear straight-through; on the rest it gathers the evidence and drafts the decision, so the handler resolves the claim in minutes rather than starting cold.

Pin it down

CRITERION 03 OF 06

03 Data residency and compliance

Residency and data protection settled in your jurisdiction, before signing.

For a regulated insurer this decides eligibility, not configuration. Where your claims data is stored and processed, and which regime governs it, are not line items you resolve after signing; they decide whether a vendor is even in the running. Name the requirement for your region before the first demo, not after.

WHAT GOOD LOOKS LIKE

- Data residency in your region offered as standard, not a paid add-on.
- A clear answer on where claims data is stored and processed.
- DPA, sub-processor, and regional compliance questions answered without a wait.

ASK THE VENDOR

Is data residency in our region standard, and where is claims data stored and processed?

REJECT IF

Residency is “on the roadmap” or priced as an upgrade.

FIG. 03 DATA RESIDENCY, IN PRACTICE

STORAGE	In your region
PROCESSING	In your region
DATA EXPORT	None without your consent

HOW MARVELX APPROACHES THIS

We treat data residency and the applicable regime as a deployment requirement, not an upsell. EU-native deployment is live today, with claims data processed in-region.

Stay in place

CRITERION 04 OF 06

04 Core-system integration

Runs alongside your claims system, not instead of it.

The strongest vendors work inside your existing claims system rather than replacing it. You keep your system of record, your reporting, and your migration budget. A rip and replace turns an AI decision into an IT programme.

WHAT GOOD LOOKS LIKE

- Claims processed inside your core system, which stays the system of record.
- Integration with the tools your workflow already runs on.
- No migration off your current platform to get started.

ASK THE VENDOR

Do you process claims inside our core system, or require us to move off it?

REJECT IF

The answer is a rip and replace.

FIG. 04 RUNS INSIDE YOUR EXISTING STACK

01	02	03
Your core claims system	MarvelX agent, in place	Back to your system of record

CONNECTS VIA

- REST & webhook APIs
- Custom connectors
- HubSpot
- IBM watsonx
- Weaviate

If your core claims, policy, or CRM system exposes an API, the agent connects to it. No migration off your platform to start.

HOW MARVELX APPROACHES THIS

We add the agent without replacing your systems. Standard API and webhook connectors plug it into the tools already around the claim, so your platform stays the system of record.

Prove it

CRITERION 05 OF 06

05 Proof over promises

Realized outcomes, not contracted targets.

The strongest evidence is a realized outcome tied to a named client and segment. A contracted target describes ambition, not track record. Ask which figures are live and which are still goals, and hold the two to different standards.

WHAT GOOD LOOKS LIKE

- A specific segment, a specific before and after, and a client who signed off.
- Figures shown on live claims, not curated samples.
- Targets clearly labelled as targets.

ASK THE VENDOR

Which figures are live with a client who signed off, and which are still goals?

REJECT IF

Contracted targets are presented as delivered results.

FIG. 05 ONE REALIZED BEFORE AND AFTER

Embedded insurance deployment, realized. Contracted motor targets are shown as targets in flight, never as results.

HOW MARVELX APPROACHES THIS

In an embedded insurance deployment, one client reached 10x claims capacity with no added headcount. Motor targets stay labelled as targets until a client signs off.

Define it

CRITERION 06 OF 06

06 Accuracy, defined

One definition, measured on live claims.

An accuracy rate means nothing without a definition and a live sample. Ask every vendor to define accuracy the same way and to show it on real claims. Speed without a defined accuracy bar only moves risk downstream.

WHAT GOOD LOOKS LIKE

- One definition of accuracy, agreed up front.
- Measured on your live claims, not curated demos.
- Reported alongside automation rate and straight-through rate.

ASK THE VENDOR

How do you define accuracy, and will you show it on our live claims?

REJECT IF

The number arrives with no definition.

FIG. 06 WHAT GOOD LOOKS LIKE IN PRODUCTION

95% decision accuracy

measured on live claims, to one agreed definition.

HOW MARVELX APPROACHES THIS

Decision accuracy is measured on live claims, underpinned by cross-document story consistency and entity-legitimacy checks inside the claim.

The Scorecard

THE SCORECARD

Score every vendor against all six.

Score each criterion from 1 to 5, weight them equally, and total out of 30. The score is a conversation, not a verdict: a vendor that fails any single criterion is rejected, not ranked below the others.

		1	2	3	4	5
01	**Explainability**					
02	**Human oversight**					
03	**Data residency and compliance**					
04	**Core-system integration**					
05	**Proof over promises**					
06	**Accuracy, defined**					
	Total	/ 30				

Give your handlers a seat at the table. They will spot the escalation gaps a slide deck hides.

Score all six

THE TAKEAWAY

Hold every vendor to all six.

A vendor that fails any one is not a trade-off you weigh. It is a rejection. Score every vendor before the demo sways anyone, and bring your handlers into the room while you do it.

Request a demo

This scorecard is yours to keep.