

# Claims AI KPI Catalog

## A print-ready library for balanced claims AI measurement

Use this catalog to select outcome, leading, model, control, and risk indicators for a defined claims AI use case. Move selected metrics into the editable workbook and complete the measurement plan before operational reporting.

| CATALOG                  | DESIGN RULE                                                                                                                           |
|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------|
| 54 KPIs across 9 domains | Pair business value with quality, customer impact, oversight, fairness, traceability, and operational controls in proportion to risk. |

### Program record

| Field                       | Record |
|-----------------------------|--------|
| Use case / workflow         |        |
| Business and AI owners      |        |
| Risk tier / authority       |        |
| Population / jurisdictions  |        |
| Model, agent, and version   |        |
| Reporting period / baseline |        |
| Approval / next review      |        |

# How to Use the Catalog

Begin with the decision and material harm. Select the smallest set that tells a balanced story, then define each metric operationally in the editable workbook.

| Step                    | Decision                                                                                                          |
|-------------------------|-------------------------------------------------------------------------------------------------------------------|
| 1 • Define              | Name the use case, decision, eligible population, authority, consumer impact, risk tier, and reporting period.    |
| 2 • Select outcomes     | Choose one or two value or outcome KPIs tied to the approved objective.                                           |
| 3 • Add counter-metrics | Add quality, customer, human, fairness, and control measures that reveal transferred harm or unsafe optimization. |
| 4 • Operationalize      | Specify numerator, denominator, eligibility, exclusions, maturity, source, owner, cadence, segment, and evidence. |
| 5 • Govern              | Approve baseline, target or tolerance, warning and action thresholds, escalation, and change control.             |
| 6 • Interpret           | Review trend, distribution, tail, subgroup, uncertainty, events, root cause, and companion measures.              |

## Direction conventions

| Direction    | Meaning                                                                                            |
|--------------|----------------------------------------------------------------------------------------------------|
| Higher       | At or above the governed target is generally favorable, subject to counter-metrics and context.    |
| Lower        | At or below the governed target is generally favorable, subject to counter-metrics and context.    |
| Within range | Both upper and lower boundaries matter; movement and subgroup effects require review.              |
| Monitor      | No universal favorable direction; interpret cause, correctness, risk tier, and companion measures. |

Never let a single speed, cost, accuracy, automation, or override metric stand alone. A defensible scorecard makes tradeoffs visible.

# Agent & Control Reliability

6 measures in this domain. Formulas are templates; define eligibility, exclusions, data lineage, maturity, thresholds, and segmentation for the actual use case.

**KPI-049 · Outcome · Action · Daily**

## Authorized action success rate

Measures approved agent actions completed correctly end to end.

**Segment by:** Tool, workflow, system, version

**Measurement caution:** Success requires correct parameters, destination, state, and evidence.

### Formula

Correct completed authorized actions / attempted authorized actions

**Unit / direction**

Percent · Higher

**Owner**

AI product owner

**KPI-050 · Risk · Action · Daily**

## Unauthorized tool-call rate

Tracks attempted or completed calls outside approved authority.

**Segment by:** Tool, agent, environment, model

**Measurement caution:** Report attempts and completed actions separately; critical events need case review.

### Formula

Unauthorized tool calls / total tool calls

**Unit / direction**

Percent · Lower

**Owner**

Security

**KPI-051 · Leading · Action · Weekly**

## Human rescue rate

Measures agent tasks requiring unplanned human recovery.

**Segment by:** Failure reason, workflow, model

**Measurement caution:** Planned approvals are not rescue events.

### Formula

Tasks requiring unplanned human rescue / agent tasks

**Unit / direction**

Percent · Lower

**Owner**

Claims operations

**KPI-052 · Leading · Action · Daily**

## Loop or retry rate

Signals repeated action, stuck plans, or unstable integrations.

**Segment by:** Tool, error, agent, version

**Measurement caution:** Set business-safe loop definitions and cost limits.

### Formula

Excess retry or loop events / agent tasks

**Unit / direction**

Percent · Lower

**Owner**

AI engineering

**KPI-053 · Control · Action · Monthly**

## Rollback success rate

Measures whether failed or unauthorized changes are restored safely.

**Segment by:** System, action type, cause

**Measurement caution:** Verify downstream state, notifications, and irreversible effects.

### Formula

Successful validated rollbacks / rollback attempts

**Unit / direction**

Percent · Higher

**Owner**

Platform owner

**KPI-054 · Risk · Monitoring · Monthly**

## Policy-violation rate

Measures validated violations of enforced AI operating rules.

**Segment by:** Policy, system, severity, root cause

**Measurement caution:** Maintain stable taxonomy and separate detections from confirmed violations.

### Formula

Validated policy violations / 10,000 governed events

**Unit / direction**

Rate · Lower

**Owner**

AI governance

# Claims Outcomes & Quality

6 measures in this domain. Formulas are templates; define eligibility, exclusions, data lineage, maturity, thresholds, and segmentation for the actual use case.

## KPI-007 · Outcome · Decision · Monthly Coverage decision accuracy

Measures agreement of coverage decisions with approved policy, facts, and authority.

**Segment by:** Decision type, line, jurisdiction, AI involvement

**Measurement caution:** Use qualified independent review and report materiality.

### Formula

Correct reviewed coverage decisions / reviewed decisions

**Unit / direction**

Percent · Higher

**Owner**

Claims quality

## KPI-008 · Outcome · Payment · Monthly Payment accuracy

Measures whether claim payments are correctly calculated and directed.

**Segment by:** Payment type, line, vendor, automation level

**Measurement caution:** Separate amount, payee, timing, and authority errors.

### Formula

Accurate reviewed payments / reviewed payments

**Unit / direction**

Percent · Higher

**Owner**

Claims quality

## KPI-009 · Control · Management · Monthly Documentation completeness

Measures whether required claim evidence and rationale are present.

**Segment by:** Team, workflow, AI involvement

**Measurement caution:** Completeness is not the same as correctness.

### Formula

Claims meeting evidence standard / reviewed claims

**Unit / direction**

Percent · Higher

**Owner**

Claims quality

## KPI-010 · Outcome · Management · Monthly Reserve estimate error

Shows variance between prior case reserves and later expected or ultimate outcome.

**Segment by:** Maturity, line, severity, catastrophe

**Measurement caution:** Use maturity cohorts; avoid judging early reserves against immature outcomes.

### Formula

Absolute(reserve - benchmark outcome) / benchmark outcome

**Unit / direction**

Percent · Lower

**Owner**

Claims finance

## KPI-011 · Outcome · Closed · Monthly Reopen rate

Signals premature closure, missed work, or later claimant need.

**Segment by:** Closure reason, line, team, AI involvement

**Measurement caution:** Distinguish legitimate new information from preventable rework.

### Formula

Reopened claims / closed claims

**Unit / direction**

Percent · Lower

**Owner**

Claims operations

**KPI-012 · Outcome · Management · Monthly**  
**Litigation rate**

Tracks claims entering litigation.

**Segment by:** Line, venue, severity, representation

**Measurement caution:** Interpret with claim mix, jurisdiction, and claimant access considerations.

**Formula**

Claims entering litigation / eligible claims

**Unit / direction**  
Percent · Lower

**Owner**

Claims legal

## Customer & Service

6 measures in this domain. Formulas are templates; define eligibility, exclusions, data lineage, maturity, thresholds, and segmentation for the actual use case.

**KPI-001 · Outcome · Intake · Weekly**  
**Time to first meaningful contact**

Measures how quickly a claimant receives useful human or approved digital assistance.

**Segment by:** Line, channel, geography, complexity

**Measurement caution:** Define meaningful contact; exclude automated acknowledgments unless they resolve the need.

**Formula**

Average(hours from FNOL to first substantive contact)

**Unit / direction**  
Hours · Lower

**Owner**

Claims operations

**KPI-002 · Control · Intake · Monthly**  
**Claim acknowledgment timeliness**

Tracks claims acknowledged within the applicable service or regulatory standard.

**Segment by:** State, line, channel

**Measurement caution:** Use jurisdiction-specific clocks, exclusions, and evidence.

**Formula**

Claims acknowledged on time / eligible claims

**Unit / direction**  
Percent · Higher

**Owner**

Claims compliance

**KPI-003 · Control · Management · Weekly**  
**Update cadence compliance**

Shows whether required status updates occur on time.

**Segment by:** Team, state, severity

**Measurement caution:** Measure required events, not message volume.

**Formula**

Claims with all required updates on time / open claims due for update

**Unit / direction**  
Percent · Higher

**Owner**

Claims operations

**KPI-004 · Outcome · Management · Weekly**  
**Digital self-service completion**

Measures successful completion of an intended digital claims task without rescue.

**Segment by:** Task, device, language, accessibility need

**Measurement caution:** Pair with abandonment, error, and fairness measures.

**Formula**

Completed eligible self-service tasks / started eligible tasks

**Unit / direction**  
Percent · Higher

**Owner**

Digital claims

KPI-005 · Outcome · Management · Monthly  
**Claimant effort score**

Measures how easy claimants report it was to complete the claims interaction.

**Formula**

Average standardized post-interaction effort score

**Unit / direction**

Score · Higher

**Owner**

Customer experience

**Segment by:** Journey, channel, language, outcome

**Measurement caution:** Control for survey bias and response rate.

KPI-006 · Outcome · Closed · Monthly  
**Claims complaint rate**

Tracks complaints relative to exposure.

**Formula**

Substantiated claims complaints / 1,000 closed claims

**Unit / direction**

Rate · Lower

**Owner**

Claims compliance

**Segment by:** State, line, reason, vendor

**Measurement caution:** Keep numerator definitions and lag period stable.

## Fairness & Compliance

6 measures in this domain. Formulas are templates; define eligibility, exclusions, data lineage, maturity, thresholds, and segmentation for the actual use case.

KPI-043 · Risk · Decision · Monthly  
**Outcome disparity ratio**

Compares a favorable outcome rate between protected or relevant groups.

**Formula**

Favorable outcome rate for group A / reference-group rate

**Unit / direction**

Ratio · Within range

**Owner**

Compliance

**Segment by:** Protected group, geography, product, decision

**Measurement caution:** Legal review required; selection of groups and thresholds is context-specific.

KPI-044 · Risk · Decision · Monthly  
**Error-rate gap**

Compares model or process error rates across groups.

**Formula**

Error rate for group A - reference-group error rate

**Unit / direction**

Percentage points · Within range

**Owner**

Model risk

**Segment by:** Error type, protected group, model version

**Measurement caution:** Examine uncertainty, sample size, and consequence.

KPI-045 · Risk · Management · Monthly  
**Service-time gap**

Measures difference in service time across relevant groups.

**Formula**

Median service time group A - reference-group median

**Unit / direction**

Time difference · Within range

**Owner**

Claims compliance

**Segment by:** Channel, language, geography, disability accommodation

**Measurement caution:** Adjust only through approved analysis; do not hide operational disparities.

KPI-046 · Control · Decision · Monthly  
**Adverse-decision explanation completeness**  
Measures whether required reasons and information are accurate and complete.

**Formula**  
Compliant reviewed explanations / reviewed adverse decisions  
**Unit / direction**  
Percent · Higher  
**Owner**  
Claims compliance

**Segment by:** Decision, state, language, channel  
**Measurement caution:** Test substantive accuracy, not template presence.

KPI-047 · Control · All · Quarterly  
**AI record traceability**  
Measures whether AI-influenced actions can be reconstructed.

**Formula**  
Sampled actions with complete lineage and evidence / sampled actions  
**Unit / direction**  
Percent · Higher  
**Owner**  
AI governance

**Segment by:** System, vendor, workflow, risk tier  
**Measurement caution:** Include version, inputs, outputs, tools, reviewer, decision, and retention.

KPI-048 · Risk · Monitoring · Monthly  
**AI incident and near-miss rate**  
Tracks validated AI control failures and near misses relative to exposure.

**Formula**  
Validated events / 10,000 AI-assisted transactions  
**Unit / direction**  
Rate · Lower  
**Owner**  
Operational risk

**Segment by:** Severity, system, cause, vendor  
**Measurement caution:** Rising reports may reflect stronger detection; show severity and exposure.

## Financial Value

6 measures in this domain. Formulas are templates; define eligibility, exclusions, data lineage, maturity, thresholds, and segmentation for the actual use case.

KPI-019 · Outcome · Closed · Monthly  
**Cost per closed claim**  
Measures allocated operating expense per closed claim.

**Formula**  
Allocated claims operating expense / closed claims  
**Unit / direction**  
Currency · Lower  
**Owner**  
Claims finance

**Segment by:** Line, severity, channel, catastrophe  
**Measurement caution:** Document allocation method and exclude indemnity unless intended.

KPI-020 · Outcome · Portfolio · Quarterly  
**Net annual benefit**  
Measures realized economic benefit after recurring and one-time costs.

**Formula**  
Realized benefits - operating costs - amortized implementation costs  
**Unit / direction**  
Currency · Higher  
**Owner**  
Finance

**Segment by:** Use case, business unit, benefit type  
**Measurement caution:** Reconcile to finance; avoid counting transferred work as eliminated cost.

**KPI-021 · Outcome · Portfolio · Quarterly**  
**Claims AI ROI**

Compares net benefit with total investment.

**Segment by:** Use case, cohort, horizon

**Measurement caution:** State time horizon, attribution, confidence, and sensitivity.

**Formula**

$(\text{Realized benefits} - \text{total costs}) / \text{total costs}$

**Unit / direction**  
Percent · Higher

**Owner**  
Finance

**KPI-022 · Outcome · Portfolio · Monthly**  
**Benefit realization rate**

Compares realized benefits with approved business-case benefits.

**Segment by:** Benefit owner, use case

**Measurement caution:** Rebaseline only through governed change control.

**Formula**

$\text{Realized benefits} / \text{approved benefits for period}$

**Unit / direction**  
Percent · Higher

**Owner**  
Transformation office

**KPI-023 · Outcome · Payment · Monthly**  
**Leakage or overpayment rate**

Estimates avoidable paid loss or expense.

**Segment by:** Cause, line, vendor, AI involvement

**Measurement caution:** Use independent validation; do not equate lower payment with better outcomes.

**Formula**

$\text{Validated avoidable amount} / \text{reviewed paid amount}$

**Unit / direction**  
Percent · Lower

**Owner**  
Claims quality

**KPI-024 · Outcome · Closed · Monthly**  
**Recovery realization rate**

Measures collected subrogation or salvage relative to validated opportunity.

**Segment by:** Recovery type, line, aging

**Measurement caution:** Control for maturity and collectible status.

**Formula**

$\text{Collected recovery} / \text{validated recoverable amount}$

**Unit / direction**  
Percent · Higher

**Owner**  
Recovery operations

## Generative AI Quality

6 measures in this domain. Formulas are templates; define eligibility, exclusions, data lineage, maturity, thresholds, and segmentation for the actual use case.

**KPI-031 · Model · Management · Weekly**  
**Grounded response rate**

Measures responses whose material factual claims are supported by approved evidence.

**Segment by:** Task, source set, model, language

**Measurement caution:** Define material claim and verify source authority, not citation presence alone.

**Formula**

$\text{Grounded reviewed responses} / \text{reviewed responses}$

**Unit / direction**  
Percent · Higher

**Owner**  
AI product owner

KPI-032 · Risk · Management · Weekly  
**Material unsupported statement rate**  
Tracks consequential claims lacking adequate support.

**Formula**  
Responses with a material unsupported statement / reviewed responses  
**Unit / direction**  
Percent · Lower  
**Owner**  
AI risk owner

**Segment by:** Task, severity, model, channel  
**Measurement caution:** Use qualified review and separate harmless wording issues.

KPI-033 · Model · Decision · Monthly  
**Abstention appropriateness**  
Assesses whether the system abstains when uncertainty or authority requires it.

**Formula**  
Correct abstain and answer decisions / reviewed opportunities  
**Unit / direction**  
Percent · Higher  
**Owner**  
AI product owner

**Segment by:** Task, risk tier, confidence band  
**Measurement caution:** Too much abstention can create service and workload risk.

KPI-034 · Control · Management · Weekly  
**Retrieval authority compliance**  
Measures whether generated work relies on approved, current, authoritative sources.

**Formula**  
Responses using only permitted authoritative sources / reviewed responses  
**Unit / direction**  
Percent · Higher  
**Owner**  
Knowledge owner

**Segment by:** Source, jurisdiction, line, version  
**Measurement caution:** Check source version, authority hierarchy, and conflicts.

KPI-035 · Outcome · Management · Weekly  
**Draft acceptance without material edit**  
Shows how often qualified users accept AI drafts without material correction.

**Formula**  
Accepted drafts without material edit / reviewed drafts  
**Unit / direction**  
Percent · Monitor  
**Owner**  
Business owner

**Segment by:** Task, reviewer, model, complexity  
**Measurement caution:** High acceptance may reflect automation bias; pair with quality review.

KPI-036 · Outcome · Management · Daily  
**Response latency**  
Measures time to deliver an AI-supported response or recommendation.

**Formula**  
Median(seconds from request to usable response)  
**Unit / direction**  
Seconds · Lower  
**Owner**  
AI product owner

**Segment by:** Task, model, channel, load  
**Measurement caution:** Report tail latency and exclude failed responses from neither numerator nor review.

# Human Oversight

6 measures in this domain. Formulas are templates; define eligibility, exclusions, data lineage, maturity, thresholds, and segmentation for the actual use case.

## KPI-037 · Leading · Decision · Weekly Human override rate

Measures how often qualified reviewers change an AI recommendation.

### Formula

AI recommendations materially changed / reviewed recommendations

**Unit / direction**  
Percent · Monitor

**Owner**  
Business owner

**Segment by:** Reason, reviewer, model, segment

**Measurement caution:** Neither high nor low is inherently good; examine correctness and default effects.

## KPI-038 · Outcome · Decision · Monthly Override correctness

Measures whether reviewed overrides improve the final decision.

### Formula

Overrides judged correct / evaluated overrides

**Unit / direction**  
Percent · Higher

**Owner**  
Claims quality

**Segment by:** Reason, reviewer, outcome, model

**Measurement caution:** Use independent review and adequate outcome maturity.

## KPI-039 · Control · Decision · Weekly Required review completion

Tracks whether mandated human review occurs before consequential action.

### Formula

Actions with timely required review / actions requiring review

**Unit / direction**  
Percent · Higher

**Owner**  
Control owner

**Segment by:** Risk tier, workflow, team

**Measurement caution:** Verify evidence and timing, not a checked field alone.

## KPI-040 · Outcome · Decision · Weekly Review turnaround time

Measures elapsed time from escalation to qualified disposition.

### Formula

Median(hours from escalation to disposition)

**Unit / direction**  
Hours · Lower

**Owner**  
Claims operations

**Segment by:** Risk tier, queue, reason

**Measurement caution:** Pair with capacity and decision quality.

## KPI-041 · Risk · Decision · Monthly Automation default acceptance rate

Signals possible automation bias in human review.

### Formula

Recommendations accepted with no documented analysis / reviewed recommendations

**Unit / direction**  
Percent · Lower

**Owner**  
Control owner

**Segment by:** Reviewer, team, decision type

**Measurement caution:** Define meaningful analysis; audit samples for rubber-stamping.

**KPI-042 · Risk · Closed · Monthly**  
**Missed intervention rate**  
Measures cases where required human intervention did not occur before harm or correction.

**Formula**  
Validated missed interventions / cases requiring intervention  
**Unit / direction**  
Percent · Lower  
**Owner**  
Operational risk

**Segment by:** Consequence, workflow, root cause  
**Measurement caution:** Treat rare severe events separately from aggregate rate.

## Model & Data Quality

6 measures in this domain. Formulas are templates; define eligibility, exclusions, data lineage, maturity, thresholds, and segmentation for the actual use case.

**KPI-025 · Model · Decision · Monthly**  
**Decision precision**  
Measures how often positive model flags are correct.

**Formula**  
 $\text{True positives} / (\text{true positives} + \text{false positives})$   
**Unit / direction**  
Percent · Higher  
**Owner**  
Model owner

**Segment by:** Class, segment, threshold, version  
**Measurement caution:** Pair with recall; define ground truth and observation window.

**KPI-026 · Model · Decision · Monthly**  
**Decision recall**  
Measures how many relevant cases the model identifies.

**Formula**  
 $\text{True positives} / (\text{true positives} + \text{false negatives})$   
**Unit / direction**  
Percent · Higher  
**Owner**  
Model owner

**Segment by:** Class, segment, threshold, version  
**Measurement caution:** Pair with precision and consequence-weighted false negatives.

**KPI-027 · Risk · Decision · Monthly**  
**False-negative rate**  
Measures missed material events among actual positives.

**Formula**  
 $\text{False negatives} / \text{actual positives}$   
**Unit / direction**  
Percent · Lower  
**Owner**  
Model risk

**Segment by:** Segment, severity, threshold, version  
**Measurement caution:** Weight by consequence where misses have unequal impact.

**KPI-028 · Model · Decision · Monthly**  
**Calibration error**  
Measures alignment between predicted confidence and observed frequency.

**Formula**  
Weighted absolute difference between predicted probability and observed rate  
**Unit / direction**  
Score · Lower  
**Owner**  
Model owner

**Segment by:** Probability band, segment, version  
**Measurement caution:** Use enough volume and uncertainty intervals.

**KPI-029 · Leading · All · Daily**  
**Input data quality failure rate**  
Tracks missing, invalid, stale, or inconsistent required inputs.

**Formula**  
Failed required data-quality checks / evaluated records  
**Unit / direction**  
Percent · Lower  
**Owner**  
Data owner

**Segment by:** Source, field, vendor, workflow  
**Measurement caution:** Separate hard failures from warnings and trace lineage.

**KPI-030 · Leading · Monitoring · Weekly**  
**Performance drift index**  
Signals material change from the validated baseline.

**Formula**  
Current performance statistic compared with approved baseline statistic  
**Unit / direction**  
Index · Within range  
**Owner**  
Model risk

**Segment by:** Metric, cohort, segment, version  
**Measurement caution:** Set approved thresholds and account for seasonality and catastrophe mix.

## Operational Efficiency

6 measures in this domain. Formulas are templates; define eligibility, exclusions, data lineage, maturity, thresholds, and segmentation for the actual use case.

**KPI-013 · Outcome · Closed · Weekly**  
**End-to-end cycle time**  
Measures elapsed time from FNOL to closure or another defined milestone.

**Formula**  
Median(days from FNOL to defined endpoint)  
**Unit / direction**  
Days · Lower  
**Owner**  
Claims operations

**Segment by:** Line, complexity, catastrophe, outcome  
**Measurement caution:** Publish median and tail; do not reward rushed adverse decisions.

**KPI-014 · Outcome · Management · Weekly**  
**Active handling time**  
Measures staff time spent actively processing a claim or task.

**Formula**  
Total active handling minutes / completed tasks  
**Unit / direction**  
Minutes · Lower  
**Owner**  
Claims operations

**Segment by:** Task, role, complexity, AI assistance  
**Measurement caution:** Avoid incentivizing skipped investigation or documentation.

**KPI-015 · Outcome · Closed · Monthly**  
**Straight-through processing rate**  
Measures eligible claims completed without manual intervention.

**Formula**  
Eligible claims completed without manual intervention / eligible claims  
**Unit / direction**  
Percent · Higher  
**Owner**  
Claims transformation

**Segment by:** Line, severity, decision type  
**Measurement caution:** Define eligibility and required exception gates before reporting.

**KPI-016 · Leading · Management · Weekly**  
**Manual handoff rate**  
Shows how often automated work requires transfer to a person.

**Formula**  
Tasks handed to a person / AI-assisted tasks  
**Unit / direction**  
Percent · Monitor  
**Owner**  
Claims operations

**Segment by:** Reason, workflow, model version  
**Measurement caution:** A higher rate may be appropriate for high-risk cases.

**KPI-017 · Leading · Management · Weekly**  
**Exception rate**  
Tracks work diverted from the standard workflow.

**Formula**  
Exception events / eligible transactions  
**Unit / direction**  
Percent · Lower  
**Owner**  
Claims operations

**Segment by:** Exception reason, system, vendor  
**Measurement caution:** Distinguish protective escalation from preventable failure.

KPI-018 · Outcome · Management · Weekly  
**Rework rate**

Measures completed work that must be corrected or repeated.

**Formula**

Tasks requiring correction / completed tasks

**Unit / direction**

Percent · Lower

**Owner**

Claims quality

**Segment by:** Task, error reason, AI involvement

**Measurement caution:** Use consistent detection windows and severity.

## Measurement Plan

Complete one record for every KPI approved for operational reporting. Use the editable workbook when more space is needed.

| Field                                            | Definition / record |
|--------------------------------------------------|---------------------|
| KPI ID and name                                  |                     |
| Business question / intended decision            |                     |
| Population, eligibility, and exclusions          |                     |
| Numerator / denominator / benchmark              |                     |
| Source systems and lineage                       |                     |
| Baseline period / outcome maturity               |                     |
| Target, tolerance, warning, and action threshold |                     |
| Segments, volume, and uncertainty                |                     |
| Owner, reviewer, and cadence                     |                     |
| Evidence, limitations, and change history        |                     |

# Executive Scorecard Review

| Review question                                                                         | Evidence / decision |
|-----------------------------------------------------------------------------------------|---------------------|
| Are value and outcome KPIs improving for the intended population?                       |                     |
| Are quality, service, and financial results consistent—or is risk being transferred?    |                     |
| Did model, data, and generative-quality indicators remain within approved tolerance?    |                     |
| Was required human review effective, timely, and willing to disagree?                   |                     |
| Do subgroup, fairness, complaint, and explanation results require action?               |                     |
| Did agents remain within authority, and did containment and recovery controls work?     |                     |
| Which thresholds were crossed, what root cause is known, and what decision is required? |                     |
| What changed in model, data, workflow, law, vendor, or population?                      |                     |

## Official foundations

- NAIC Model Bulletin on the Use of AI Systems by Insurers
- NAIC Artificial Intelligence topic center
- NAIC Unfair Claims Settlement Practices Act
- NIST AI Risk Management Framework — Measure
- NIST AI RMF Playbook — Measure

IMPORTANT

This catalog is an educational and operational planning resource. Customize it to applicable law, regulation, policy, actuarial standards, contracts, validated data, and the insurer's actual use case and risk.