A Market-Conduct-Safe, Evidence-Anchored Decision Packet for AI-Assisted P&C Underwriting

Design principles for auditability, adverse-action defensibility, and controlled discretion

Upendra Belhe, Ph.D.
Belhe Analytics Advisory
January 2026

Abstract

AI-assisted underwriting in regulated P&C insurance is constrained not by model capability, but by the inability to produce durable decision evidence that withstands market conduct review. Regulators audit files, not models. This paper proposes a Decision Packet: a standardized, auditable, replayable record binding underwriting outcomes to contemporaneous evidence, governing constraints, decision authority, and alternatives considered. We define a protocol comprising a message envelope, evidence bundle contract, decision graph, and compliance replay manifest. Two worked examples (Personal Lines Auto and Commercial Workers' Compensation) demonstrate implementation. The proposed construct is buildable today and enables scalable, defensible AI-assisted underwriting in regulated environments.

1. Problem Statement

P&C underwriting decisions persist beyond model lifecycles. Policies renew for years, losses emerge later, disputes arise later still, and regulators review historical files after systems and teams have changed. Traditional AI deployments provide execution traces and feature attributions that suffice for engineering diagnostics but fail to satisfy regulatory, legal, and audit expectations. This gap manifests as market conduct risk, adverse-action defensibility risk, and erosion of customer trust. A new evidentiary primitive is required: the Underwriting Decision Packet.

2. Design Goals

A Decision Packet must enable: (i) reconstruction of the decision as made; (ii) proof of active constraints (filings, guidelines, underwriting authority); (iii) provenance and reliability of material signals; (iv) controlled discretion and override accountability; (v) examiner-grade narratives without post-hoc storytelling; and (vi) scalable standardization across products and lines of business.

3. Core Principle

Underwriting AI must produce evidence, not just outputs. Engineering-grade artifacts describe execution. Market-conduct-grade artifacts must describe decision state, constraints in force, authority, evidence, and alternatives considered.

4. Threat Model for Market Conduct Failures

This protocol targets common underwriting failure modes: (1) silent appetite drift (routing/triage changes that reshape the book without an explicit appetite change); (2) unprovable adverse actions (declines or restrictive terms lacking contemporaneous evidence and rule references); (3) untracked discretion and overrides (human or system deviations without bounded authority and rationale); (4) version ambiguity (unclear model/ruleset/guideline versions at decision time); and (5) evidence without provenance (external signals used without durable pointers, timestamps, and reliability grading).

5. The Underwriting Decision Packet Construct

A Decision Packet is an immutable, structured record comprising: (1) a Message Envelope (purpose, scope, versions, authority); (2) an Evidence Bundle (provenance, reliability, minimal excerpts); (3) a Decision Graph (branch logic, constraints, counterfactuals); (4) an Outcome Record (action and effective terms); and (5) a Compliance Replay Manifest (deterministic reconstruction recipe).

5.1 Message Envelope Schema (minimal)

```
{
 "envelope_version": "1.0",
 "submission_id": "SUB-2026-000812",
 "risk_id": "RISK-7F2C1",
 "line_of_business": "PersonalAuto",
 "jurisdiction": "TX",
 "channel": "DigitalDirect",
 "timestamp_utc": "2026-01-24T02:15:22Z",
 "purpose": "UnderwritingEligibilityAndTerms",
 "privacy_scope": ["PII_MINIMIZED", "REGULATED_SIGNALS_OK
      "],
 "guideline_version": "UWGL-PA-TX-2025.12",
 "filing_id": "SERFF-TX-PA-2025-09-44110",
 "ruleset_version": "RULES-PA-TX-4.2",
 "model_version": "MODEL-UW-TRIAGE-2.1.4",
 "authority": {
 "actor_type": "SYSTEM",
 "actor_id": "UDA",
 "authority_tier": "TIER_1",
 "max_surcharge_pct": 15,
 "must_refer_conditions": ["DUI_WITHIN_36M", "LAPSE_GT_30D
      "]
 },
 "audit": {
 "correlation_id": "CORR-91C0",
 "decision_packet_id": "DP-2026-01-000812",
 "immutability": "WORM_REQUIRED"
 }
}
```

6. Evidence Bundle Contract

Evidence is not merely a feature value; it is a claim about reality, bound to provenance, retrieval time, reliability, and relevance to an underwriting constraint. Evidence below a defined reliability threshold should not drive adverse outcomes without human review.

6.1 Evidence Item Schema (example)

```
{
 "evidence_id": "EVD-00112",
 "type": "MVR",
 "source": {
 "provider": "StateDMV-Partner",
 "artifact_uri": "uw://evidence/SUB-2026-000812/mvr.json",
 "hash": "sha256:9b0c...",
```

```
 "retrieved_at_utc": "2026-01-24T02:16:10Z"
 },
 "minimal_excerpt": {
 "violations_36m": 2,
 "major_violation": false,
 "violation_types": ["Speeding_15Over", "Speeding_10Over"]
 },
 "reliability_grade": "A",
 "why_it_matters": "Violations exceed standard-tier
     threshold in UWGL-PA-TX-2025.12.",
 "constraints_touched": ["RULES-PA-TX-4.2#TIERING.
     VIOLATIONS"],
 "countervailing_factors": ["No prior claims in CLUE;
     continuous coverage verified."],
 "pii_classification": "MEDIUM"
}
```


7. Decision Graph and Counterfactuals

The Decision Graph records which constraints were evaluated, which branch fired, what evidence supported the branch, and which near alternatives were considered and rejected. Counterfactuals are required for examiner-grade defensibility: why a less adverse option was not selected.

7.1 Decision Node Schema (example)


```
{
 "node_id": "N-03",
 "node_type": "RULE_EVAL",
 "rule_ref": "RULES-PA-TX-4.2#TIERING.VIOLATIONS",
 "inputs": ["EVD-00112"],
 "evaluated_state": {"violations_36m": 2},
 "result": "APPROVE_WITH_SURCHARGE",
 "rationale": "Violation count exceeds standard tier;
     surcharge allowed within authority.",
 "alternatives_considered": [
 {"action": "APPROVE_STANDARD", "rejected_because": "tier
     threshold exceeded"},
 {"action": "REFER_TO_HUMAN", "rejected_because": "no must-
     refer condition triggered"}
 ]
}
```


8. Worked Example 1: Personal Lines Auto

Scenario. A Texas personal auto submission is evaluated for eligibility and pricing. The system proposes approval with a bounded surcharge based on driving history.

Evidence Items. (E1) MVR indicating two minor speeding violations within 36 months (grade A). (E2) CLUE Auto loss history: no prior claims (grade A). (E3) Prior insurance verification: continuous coverage, no lapse (grade B/A depending on provider). (E4) Application attestation and signed disclosures (grade A).

Decision Graph (summary). N-01 completeness checks pass. N-02 eligibility constraints pass (no DUI, no major violations). N-03 tiering rule fires based on MVR: surcharge required for standard tier. N-04 authority check confirms Tier-1 can apply up to +15% surcharge. Counterfactuals. Standard rate rejected due to tier threshold. Referral rejected because must-refer condition not met.

Outcome. APPROVE_WITH_SURCHARGE. Decision Packet seals: envelope (purpose/scope/versions/authority), evidence bundle (provenance + hashes), decision graph nodes and counterfactuals, and replay manifest.

9. Worked Example 2: California Workers' Compensation

Scenario. A California light manufacturing firm submits a WC application. The system evaluates class codes, payroll exposure, and loss experience. It proposes referral due to elevated loss frequency and inspection findings.

Evidence Items. (E1) ACORD application with payroll and WC class codes (grade A). (E2) Prior carrier loss runs showing three lost-time claims within five years (grade A). (E3) Safety questionnaire responses (grade B). (E4) Third-party inspection report noting machine guarding deficiencies (grade B/A).

Decision Graph (summary). N-01 validation passes. N-02 appetite rule triggers referral based on loss frequency threshold for class code group. N-03 inspection evidence confirms elevated hazard controls gap. N-04 authority check: system lacks binding authority for this referral condition. Counterfactuals. Standard approval rejected (appetite violation). Decline rejected (eligible with risk improvement conditions).

Outcome. REFER_TO_HUMAN. Human underwriter applies a risk improvement endorsement (guarding remediation) and binds subject to conditions. Decision Packet records human authority, conditions, and the exact evidence/rules that triggered referral.

10. Compliance Replay

Compliance Replay deterministically reconstructs the underwriting decision using archived evidence artifacts and recorded versions of rulesets, guidelines, and models. Replay forbids external calls and seeds any nondeterministic component. This enables examiner-grade reproduction months or years later.

10.1 Replay Manifest (minimal)


```
{
 "decision_packet_id": "DP-2026-01-000812",
 "replay_version": "1.0",
 "inputs": {
 "envelope_hash": "sha256:3d12...",
 "evidence_hashes": ["sha256:9b0c...", "sha256:1aa7..."]
 },
 "execution": {
 "ruleset_version": "RULES-PA-TX-4.2",
 "guideline_version": "UWGL-PA-TX-2025.12",
 "model_version": "MODEL-UW-TRIAGE-2.1.4"
 },
 "determinism": {
 "random_seed": 1729,
 "external_calls": "DISALLOWED_DURING_REPLAY"
 },
 "outputs": {
 "decision_graph_hash": "sha256:77af...",
 "outcome_hash": "sha256:0c91..."
 }
}
```


11. Implementation Notes

Decision Packets should be stored in immutable (WORM) storage with indexed registries. The policy administration system should store a stable pointer to the Decision Packet ID for renewals and endorsements. Schema governance should treat the Decision Packet as a governed data product, with semantic versioning for envelope, evidence, and decision graph schemas.

12. Conclusion

AI-assisted underwriting scales only when underwriting files can explain themselves. Market conduct is a file problem, not a model problem. The Decision

Packet converts AI outputs into examiner-grade evidence: provenance-aware, constraint-bound, authority-limited, and replayable. This is the foundation for deploying agentic underwriting in regulated P&C environments.