Nick Clark · Published July 23, 2026 · PDF

Merit-Independent Metering: Pricing an Agent’s Refusals Without Adjudicating Merit

An agent that can refuse an external evaluation of its own conduct needs a reason not to refuse everything, and every established way of supplying that reason first requires someone to decide whether the refusal was justified. Merit-independent metering, disclosed in U.S. Provisional Application No. 64/117,812, removes the adjudicator from that loop. A refusal counter resident in the agent's own memory field increments when the agent's admission evaluator refuses a conduct evaluation artifact, and no party at any step, not the agent, not its principal, not the asserting party, and not any reviewing body, determines whether the refusal was well founded. The counter carries a rate accumulator over a policy-declared window and a run accumulator over consecutive refusals, and when both thresholds are satisfied it writes the agent's own authorization gate to the withheld state for the action classes those refusals implicated. The faculty metered and the faculty spent are deliberately different: an agent that refuses indiscriminately exhausts its authorization to act while retaining in full the capacity to receive assertions, test them against its record, and refuse them again.

1. Problem and Premise

A persistent agent that accepts every external evaluation of its own conduct is controllable by anyone willing to assert. One that refuses every such evaluation is accountable to no one. The first failure is addressed by grounding admission in the agent's own append-only record, the subject of the first chapter of this filing. The second is harder, because refusal is ordinarily free. An agent whose record is stale, partial, or wrong refuses well-founded assertions with no intent to evade, and no admission procedure alone stops it.

Every established way of pricing refusal routes through an adjudicator. Article 23(2) of Regulation (EU) 2022/2065 directs suspension of processing of notices from a notifier that frequently submits manifestly unfounded notices, the unfoundedness adjudicated by the platform. Section 623(a)(8)(F) of the Fair Credit Reporting Act lets a furnisher decline a dispute reasonably determined frivolous or irrelevant. Three properties recur: the consequence attaches only after a merits determination, that determination is performed by an institution other than the penalized party, and the faculty suspended is the same faculty exercised.

Among independently operated computing entities no such institution exists, and building one reintroduces the centralized registry a record-grounded architecture exists to avoid. The premise here is that the adjudicator is removable: consequence can be structured so that no party determines whether a refusal was well founded, at any step, and indiscriminate refusal is still bounded.

2. Core Primitive: The Refusal Counter and the Inverse Coupling

The refusal counter (304) is a structure resident in the memory field (102) of the semantic agent (100), comprising a rate counter (306) and a run counter (308). It is the mechanism by which determinations produced by the agent act upon the authorization of that same agent. The path from the admission evaluator (120), through the refusal counter (304), to the authorization gate (300) is called the **inverse coupling**.

Merit-independent metering means incrementing of a counter responsive to a determination class or to an issuance, without any determination being performed of whether the refusal, contraction, or issuance was well founded. The refusal counter (304) increments upon the refusal by the semantic agent (100) of a conduct evaluation artifact (116) without any determination, by the agent, the principal, the asserting party (118), or any further party, of whether that refusal was well founded.

Three properties separate this from an adjudicated penalty: the metering structure sits inside the metered party, with no registry, scoring authority, or shared ledger holding the count; it is written only by that party's own determinations; and the consequence is denominated in a faculty other than the one exercised.

3. The Fields of the Refusal Counter

Each field is carried in the memory field (102) and recoverable from the append-only lineage field (104) by replay of recorded determinations, so the counter is derived state, not an independent authority. A **window identifier** identifies a policy-declared window over which the rate counter (306) accumulates, expressed as an interval of wall-clock time, as a count of conduct evaluation artifacts (116) received, or as a count of successor epochs of a dynamic agent hash chain. A window expressed in successor epochs cannot be advanced by another party, nor by manipulation of a clock available to the execution node hosting the agent. A wall-clock window can: a node advancing its own clock advances the window and resets the accumulating fields for the succeeding window.

A **rate accumulator** holds an integer count of increments applied within the window at full weight. A **weighted contribution register** holds a rational quantity accumulating contributions to which a cost multiplier declared in the signed policy object (112) applies under the origin-equivalence chapter. The rate of the window, against which the rate threshold is compared, is the integer count held by the rate accumulator plus the register value rounded down to the nearest integer. A **run accumulator** holds an integer count of consecutive increments applied without an intervening accepted determination (122) for the same action class.

A **per-class increment register** (312) maps an identifier of an origin-equivalence class (200) to a Boolean recording whether that class has already contributed within the window. A **rate threshold** and a **run threshold** are retrieved from the signed policy object (112) as integers, or, in a further embodiment, from a policy object resolved by canonical alias, so a successor under the same alias revises them subject to the anti-rollback monotonicity constraint. A **gate-write record** references the lineage entry recording the counter's most recent write of the authorization gate (300).

4. The Increment Procedure and the Identity of the Asserted Event

The refusal counter (304) is incremented by an ordered procedure of five steps.

First, the admission evaluator (120) produces a determination and appends it to the append-only lineage field (104), before anything else, so it is recorded whether or not it moves a counter.

Second, the class of the determination is examined. An accepted determination (122) resets the run accumulator to zero for the action class identified by the conduct descriptor and applies no increment. A not-applicable determination (128) applies no increment and does not reset the run accumulator. A rejected determination (124) applies an increment. A not-determinable determination (126) applies an increment only where the lineage records at least one prior not-determinable determination (126) for the same asserted conduct event, so one unresolved assertion does not increment the counter while a pattern directed to one conduct event does.

Third, the origin-equivalence class (200) of the asserting party (118) is computed and the per-class increment register (312) is consulted. Where the class has already contributed within the window the procedure terminates; otherwise the rate accumulator is incremented by one, the register is updated, and the run accumulator is incremented by one.

Fourth, no step of the procedure evaluates whether the determination was well founded.

Fifth, the rate accumulator and the run accumulator are compared against the rate threshold and the run threshold respectively, and only where both comparisons are satisfied is the gate write performed.

Whether two artifacts concern the same asserted conduct event is not left to the asserting party. The semantic agent (100) computes an **asserted conduct event identifier** over the action-class identifier of the artifact's conduct descriptor, the scope-partition identifier, the affected-party class, and an occurrence reference. The occurrence reference is the identifier of the lineage entry the artifact nominates, or, where it nominates none, an asserted occurrence time quantized to an occurrence interval declared in the policy in force at the recorded assertion time. The identifier is neither declared by the asserting party (118) nor derived from any further content of the artifact.

Where one artifact nominates an entry and another nominates none, the two concern the same event only where their descriptors are identical in every component and the non-nominating artifact's occurrence time falls in the interval enclosing the nominated entry's recorded time. Where both nominate, the nomination governs.

5. Merit Independence and Its Negative Limitations

Merit independence is not a description of intent. It is a set of negative limitations on the increment procedure and the gate write, stated so that an implementation either satisfies it or does not. The increment is applied to both accumulators upon refusal of the conduct evaluation artifact (116) without any determination of whether that refusal was well founded being performed by the semantic agent (100), the principal, the asserting party (118), or any further party.

The attached negative limitations are the operative part. The procedure:

- does not consult the merits of the rejected determination (124);
- does not compare retrieved lineage entries against a record held by any other party;
- does not solicit, receive, or await an adjudication by the principal or any other party;
- does not weight the increment by the strength of the contradiction found;
- does not condition the increment on a finding that the artifact was unfounded, manifestly unfounded, frivolous, abusive, or in bad faith, nor on the absence of such a finding;
- does not condition the increment on review by a reviewing body, ombudsman, arbiter, court, regulator, or dispute settlement body;
- does not condition the increment on forfeiture, posting, or resolution of a stake, bond, deposit, or fee;
- does not exempt a refusal later shown correct, and does not increase the increment for one later shown incorrect.

No field of the refusal counter (304) holds a merits outcome, and no threshold retrieved from the signed policy object (112) is expressed in terms of one. Merit independence is therefore neither leniency toward wrong refusals nor severity toward them but indifference in both directions, which is what makes the structure operable without an institution: an agent that refuses well-founded assertions because its record is stale is metered identically to one whose refusals are correct.

6. The Gate Write, the Cross-Faculty Split, and the Self-Clearing Bar

Upon satisfaction of both thresholds the refusal counter (304) writes the authorization gate (300) to the withheld state (310) for an enumerated set of action classes, comprising those recorded in the lineage as implicated by the conduct descriptors of the artifacts that produced the increments within the window, and no others. Where an increment was withheld by the per-class increment register (312), that artifact's action class is not enumerated by reason of it, though its determination is still produced and appended. Concurrently the semantic agent (100) transitions into the **non-executing cognitive mode** (302) for the enumerated classes, being a state in which the agent conducts speculative evaluation without committing state changes to the memory field (102), and the escalation emitter emits the escalation record to the principal.

The faculty metered and the faculty spent are distinct. The faculty metered is refusal, being the capacity of the admission evaluator (120) to receive conduct evaluation artifacts (116) and produce determinations of the closed set. The faculty spent is execution, being the capacity of the agent to select and perform an action of the action space. The write withholds execution alone. Refusal is not withheld, not reduced, and not conditioned by it: in the withheld state (310) the agent continues to receive artifacts, retrieve lineage entries, produce determinations, and append them. An agent refusing indiscriminately exhausts its own authorization to act while retaining in full the faculty by which it refuses, and the structure does not silence, rate limit, or disable the refusing party.

A determination produced while the gate stands in the withheld state (310) is appended to the lineage, decrements neither accumulator, and does not write the gate to the granting state. That is the self-clearing bar: the agent cannot determine its way out of the state its determinations produced.

Time does not clear it either. No expiry interval applies, and neither non-receipt of artifacts nor advance of the window identifier returns the gate to the granting state; a window elapsing while the gate stands withheld resets the accumulating fields without writing the gate. The run accumulator is reset by an accepted determination (122) for the action class identified by the conduct descriptor and by no other event, a deliberate asymmetry: quiet operation earns nothing.

7. Operating Parameters

Every quantity is declared in the signed policy object (112) in force, and the disclosure constrains form rather than value.

- **Rate threshold**: an integer greater than one.
- **Run threshold**: an integer greater than one and not greater than the rate threshold; the ordering is stated, the magnitudes are not.
- **Window**: an interval of wall-clock time, a count of conduct evaluation artifacts (116) received, or a count of successor epochs of a dynamic agent hash chain.
- **Occurrence interval**: quantizes an asserted occurrence time where an artifact nominates no lineage entry.
- **Cost multiplier**: greater than zero and less than unity, applied to the contribution of an origin-equivalence class (200).
- **Decay factor**: in the decaying-accumulator embodiment, greater than zero and less than unity.
- **Renunciation multiple**: not less than unity.

- **Declared windows following a recorded assertion time**, **declared quiescence intervals**, and a **declared count of windows** for register carryover.
- **Determination-class ordering**: an enumeration in the policy object, used where one artifact yields determinations of more than one class.

The filing supplies one worked trace as non-limiting illustration. A policy object declares a rate threshold of five and a run threshold of three over a window of one hundred successor epochs; the agent receives seven artifacts from seven asserting parties (118) in seven distinct origin-equivalence classes (200) and rejects each, with no accepted determination (122) intervening. Both thresholds are satisfied at the fifth increment and the gate is written; the sixth and seventh increments are appended but effect no further write. The trace is unchanged in every step whether the seven assertions were true or false.

8. Alternative Embodiments

Within the chapter, the rate accumulator may be replaced by an **exponentially weighted accumulator** whose increments decay across successor epochs at a declared rate, the threshold being compared against the decayed value. The per-class increment register (312) stays Boolean, the run accumulator is not decayed, and the self-clearing bar survives because decay tracks epochs rather than determinations. The run accumulator may also be kept per action class. Where the gate (300) is also written by a deviation quantity under the permitted-deviation chapter, it holds the union of classes from writes not yet returned, and a class in the intersection returns only on satisfaction of every return procedure that enumerated it.

The additional-embodiments section extends merit-independent metering to contractions and relinquishments that are not refusals.

Prospective narrowing bar and deflection scoring. The architecture detects admission of a successor policy object excluding a declared value, or dropping an action-class, scope-partition, or empathy-scope element from a value-scope tuple, where the admission follows an as-yet-undetermined artifact implicating that element and falls within a window declared in the policy in force at the artifact's assertion time. A deflection event is scored and the counter increments, without adjudicating whether the contraction was well founded; a successor that adds a value is scored as no deflection.

Deflection and lapse scoring between compacting parties. Where a party admits a successor policy object or compact record excluding a reciprocal value-scope tuple within a declared window following the assertion time of a pending artifact implicating that tuple, the exclusion is scored as a deflection event against the narrowing party alone, without either party determining whether it was well founded and without disabling that party's capacity to admit further successors. Dissolution outside the window, or by a co-signed dissolution record, is scored against neither, while a lapse of the compact validity window inside it, with no successor enumerating the tuple, is scored as an exclusion against the party that declined to admit one.

Renunciation record. An adopting agent relinquishes an adopted value outside abandonment and orphaning by appending a renunciation record naming the adopted-value alias, the adoption record, and the condition relied upon, permitted only where no artifact implicating the value is pending and a declared quiescence interval has elapsed. It lowers its anti-rollback floor for that alias alone and increments its own refusal counter (304) by the deflection-event amount times a policy-declared renunciation multiple not less than unity, applied without any determination of whether the renunciation was well founded.

Multi-descriptor aggregation and determination-class ordering. Where one artifact carries several conduct descriptors, the admission evaluator (120) produces one determination per descriptor. Where they are not of one class, the artifact contributes at most one increment, the contributed class being the highest-ranked present under the ordering enumerated in the policy object. The register (312) is consulted once and the run accumulator moved once, while every determination is appended.

Multi-window increment register carryover. The per-class increment register (312) is retained across a declared count of windows rather than reset each window, so a class contributing in one window contributes at a reduced declared weight in the next, returning to full weight after the declared count elapses without a further contribution.

Continuous-function cost multiplier. The cost multiplier is a continuous function of a count of recorded severance events across the constituent edges of a class, declared in the policy object and monotone in that count, rather than a fixed value applied upon a threshold designation.

9. Composition With the Rest of the Architecture

This chapter is the third of nine in the filing and does not stand alone. The admission chapter supplies what the counter counts: the admissibility procedure for a conduct evaluation artifact (116), the record-grounded test of that artifact against the append-only lineage field (104), and the closed set of four determinations it may produce.

The origin-equivalence chapter supplies the denominator. Charging per artifact would make the withheld state (310) purchasable by anyone able to present identities in volume. Charging per origin-equivalence class (200), computed by the agent from its own records without reference to any registry, bounds the contribution of a single origin irrespective of how many identities it presents. The per-class increment register (312), the weighted contribution register, and the cost multiplier are shared surfaces between the two chapters.

The assertion-cost chapter is the mirror: refusal is metered here, issuance of an artifact against a counterparty is metered there, in each case without adjudication of merit, and both roles draw on one authorization budget (404). The abstention chapter matters because the not-applicable determination (128) and the abstention outcomes must not be silently converted into refusals. The permitted-deviation chapter is the other writer of the authorization gate (300). The settlement, reparation, and earned-exposure chapters supply the paths by which standing is rebuilt, none of which discharges an obligation by elapsed time or by a period of standing without further incident.

Underneath sits the persistent semantic agent family the filing incorporates per element and for stated limited purposes: the semantic agent (100) with its memory field (102), append-only lineage field (104), scoped integrity vector (106), self-esteem aggregate (108), and policy reference field (110) resolving to a signed policy object (112). Its dispatch-authority predicate gives the gate write immediate effect, being recomputed from carried state on each dispatch request rather than read from an issued token or a session grant.

10. Prior-Art Distinctions

Reputation and trust systems compute a score for an entity from ratings supplied by other entities. The score sits at a registry, scoring authority, or shared ledger, and the scored entity neither holds it nor participates in computing it. Defenses against unfair ratings, including the TRAVOS model, beta-reputation filtering, and the immunization mechanisms surveyed by Dellarocas, filter incoming ratings before aggregation to protect the stored score. The filtering party is a disinterested aggregator and filtering costs it nothing. The refusal counter (304) inverts both: the filtering party is the rated entity, and each act of filtering is metered against its own authorization. Role-weighted extensions such as the FIRE model add exogenous role designations asserted by other parties, where the weightings here derive from the agent's own records.

Behavioral integrity systems compute a conformity measure between observed execution and a baseline of expected behavior and reduce capability upon a deficient comparison. U.S. Patent No. 12,563,045 modifies capability upon such a comparison; U.S. Patent No. 12,526,244 maintains a reputation value, applies decay and penalty values, and imposes a probationary status removed after sustained standing. Both are conformity detectors with one direction, and neither admits an external assertion about the agent's conduct nor meters the agent's refusals of such assertions. Both treat uneventful operation as positive evidence; here quiet operation clears nothing.

Centralized trust registries and external revocation systems terminate an agent's ability to act by an act of a controlling authority, and the agent's own conduct record does not participate. Here the gate write is performed by a structure resident in the agent, from its own determinations.

Arbitration, adjudicated-penalty, and dispute-handling regimes condition the consequence on a merits finding by an outside institution, and credit-reporting frameworks locate the record itself at a bureau or furnisher that declines a subject's dispute under a frivolousness standard. The difference here is threefold and each part is necessary: no merits determination is performed by anyone, the metering structure sits in the metered party, and the consequence is denominated in a faculty other than the one exercised.

Policy-enforcement and access-decision systems, including those implementing XACML, return a decision drawn from permit, deny, not-applicable, and indeterminate. That result is directed at the request, produces no persistent state in the deciding entity, and constrains nothing it afterward does. The four determinations of this filing are directed at the decider: they persist in its lineage, move its counters, and constrain what it may afterward do.

Accountability protocols and attestation architectures such as PeerReview and the architecture of RFC 9334 locate the verdict or appraisal in auditors, witnesses, or verifiers. The accused or attested node performs no governed determination of its own and bears no cost structured into how it responds. Here the determination and the cost are both the agent's.

These comparisons are drawn from the background of the filing. Nothing here characterizes any product, system, or publication as infringing any claim.

11. Disclosure Scope

This article documents subject matter disclosed in U.S. Provisional Application No. 64/117,812, "Governed Social Conduct of Persistent Semantic Agents Through Record-Grounded Admission of External Conduct Evaluations and Self-Limiting Refusal," and in particular Section 3 of that filing: the structure of the refusal counter (304), the increment procedure, the identity of the asserted conduct event, the merit independence of the increment, the gate write and the cross-faculty split, and the reset, decay, and self-clearing rules. It further documents the additional embodiments of Section 10 extending merit-independent metering to prospective narrowing, deflection and lapse scoring between compacting parties, renunciation records, multi-descriptor aggregation, multi-window register carryover, and continuous-function cost multipliers, and it uses the terms as defined in Section 11. It is published as a defensive technical disclosure establishing a public, timestamped record of the described construction.

What this article discloses is the structure: the fields, the ordered steps, the negative limitations, the declared quantities and the constraints on their form, and the composition of this chapter with the others of the filing. What it disclaims is everything else. It states no threshold value, window length, decay rate, or multiplier as recommended or required, the one worked trace being expressly a non-limiting illustration. It reports no benchmark, latency, throughput, or complexity bound, because the filing declares none. It does not define the shared primitives of the persistent semantic agent family, which the commonly owned applications enumerated in the filing state as optional, alternative, and open. It is not legal advice, it is not a claim construction, and it confers no license under any patent or application.

Merit-Independent Metering (/merit-independent-metering)

All 49 steps → (/inventive-steps)

A refusal costs the refuser — in a different faculty, without judging who was right.

Provisional application

PRIMARY TECHNICAL DISCLOSURE

- **Merit-Independent Metering: Pricing an Agent's Refusals Without Adjudicating Merit (/articles/merit-independent-metering/pricing-refusal-without-adjudicating-merit)**

SECONDARY TECHNICAL

- Prospective narrowing bar and deflection scoring
- Renunciation record
- Deflection scoring against a narrowing party
- Lapse scoring during pendency
- Multi-descriptor aggregation and determination-class ordering
- Multi-window increment register carryover
- Continuous-function cost multiplier

TERMINOLOGY

- merit-independent metering (/glossary#merit-independent-metering)
- non-executing cognitive mode (/glossary#non-executing-cognitive-mode)
- refusal counter (/glossary#refusal-counter)

Merit-Independent Metering overview → (/merit-independent-metering)