

The Three-Layer Separation

Status: Architectural invariant. Hard rule.

Owner: Randall Gossett

Version: v3.1.2 (canonical)

Last updated: 2026-05-08

TL;DR

The Extropy Engine has three layers. They must stay separate. Confusing them is how every reputation system in history has eaten itself.

Layer	Visible to	Currency	Purpose
1. User-facing	Everyone	Discounts, savings, gamified feedback	Make participation feel good and pay off in dollars
2. Merchant-facing	Businesses	Better POS, customer pipeline, operational signal	Make merchants want in, without charging them SaaS
3. Engine	Validators, sensors, the math	XP, CT, CAT, IT, DT, EP	Actually measure entropy reduction. Never user-visible as a score.

Hard rule: Layer 1 never exposes raw XP as a number a user can target. Layer 3 never gets simplified into a leaderboard. The gamification on Layer 1 is a deliberate decoy. The real scoring function lives where users can't farm it.

Why this separation exists

Every gamified scoring system in history has been gamed.

- Credit scores became things people optimize for instead of actual creditworthiness.
- Klout became a parody of itself.
- China's social credit became coercive.
- FICO is a number people target, not a true measure.
- Steps, streaks, productivity dashboards: people fake the number, abandon the underlying behavior.

This is Goodhart's Law: when a measure becomes a target, it stops being a good measure.

The standard response to Goodhart is to hide the metric. That's what we do, with one twist: we don't hide it because we're being secretive. We hide it because the *user-visible* gamification is intentionally a different metric than the *system-level* scoring function.

The user can farm the visible one all day. It pays out in fun, dopamine, badges, streaks. It does not pay out in actual EP. Actual EP comes from the underlying entropy-reduction signature, which is computed on Layer 3 from validator-witnessed data. The user can't push that lever directly because they don't know which lever it is, and even if they did, validators would catch the manipulation.

This is structurally different from "hidden scoring." It's **deliberate metric divergence** between what the user sees and what the system rewards. The two are correlated by design. They are not the same.

Layer 1: User-facing surface

What the user sees

- Discounts at participating merchants
- Money saved this month, this year
- Gamified feedback: streaks, completed loops, household achievements, "you helped your community offset 2.3 tons of food waste"
- A character sheet they curate. They hold the keys. They choose what to share with whom.
- Levels, badges, narrative hooks. Whatever pulls people in.

What the user does NOT see

- Raw XP numbers
- Their entropy-reduction coefficient
- Their rarity multiplier per domain
- Any direct lever that maps 1:1 to EP minting

Why

If a user can see "I earned 47 XP for taking out the recycling," they will start taking out the recycling 30 times a day to grind. The XP math has to be opaque enough that the user can only optimize by *actually doing the thing the system is trying to reward*.

Practical translation: the user sees "you saved \$42 this month by cooking at home 3+ nights a week." They don't see the formula behind it. The formula could change. The savings stay legible.

The character sheet metaphor

The user holds a self-curated record of their habits and contributions. Validators and sensors prevent fabrication. They can choose what to display, but they can't forge what's there.

You hold the pen and the eraser. Reality holds the dice.

This is the elevator pitch. It's the difference between this and a social credit score: in a social credit system, the state holds the pen. In ours, the user does. The state, employer, or government can't write to your sheet. Only validated reality can.

Allowable gamification on Layer 1

- Streaks (decay-based, so they reset and don't compound forever)
- Levels (cosmetic, not consequential beyond a flavor identity)
- Household / community challenges (collective, not zero-sum)
- "You're in the top X% of households doing Y" (relative, contextual)
- Narrative achievements ("First Loop Closed," "Foundational Validator")
- Visualization of entropy patterns (artistic, not numeric)

Forbidden on Layer 1

- Public XP leaderboards
- Raw EP balance as a status symbol
- Anything that creates social pressure to over-report
- Anything that ties self-disclosure volume to reward magnitude
- Anything that lets a user reverse-engineer the formula by varying inputs

Layer 2: Merchant-facing surface

What the merchant sees

- A free or near-free POS that runs all standard payment flows
- A customer pipeline: people on the network preferentially shop at network merchants
- Better operational signal than KPIs (entropy-reduction patterns reveal what's actually working in their business)
- Standard merchant-services infrastructure (1.8% rates instead of 2.5–3%)
- Optional DFAO node registration for deeper benefits
- Self-reporting tools for inventory, sourcing, supply chain

What the merchant does NOT see

- Individual user XP balances
- Individual user character sheets (unless the user chooses to share)
- The full Layer 3 math
- Any way to manipulate a customer's score

Why merchants opt in

Merchants don't opt in because they believe in entropy reduction. They opt in because:

1. The POS is free or cheaper than what they're using.
2. Customers are walking in asking which businesses are on the network.
3. The data they get back is operationally better than legacy KPIs.
4. The customers attracted by the network are higher-quality (more loyal, more habitual, lower acquisition cost).

The entropy-reduction layer is, from their perspective, an invisible substrate that happens to make the business case work. They don't have to understand it. They just have to use it.

Monetization (without charging the user, without paywalling the POS)

- Capture the standard merchant-services fee (the 1.8% they were already paying someone else)
- DFAO node registration fees at scale
- Treasury yield on EP float
- Premium analytics for businesses that want deeper signal
- Multinational nodes paying for specialized integrations

Don't charge SaaS. Don't paywall the POS. Don't bill the user. The two flywheel rules.

Layer 3: The engine

What lives here

- The XP formula: $XP = R \times F \times \Delta S \times (w \cdot E) \times \log(1/T_s)$
- The CT formula: $CT = C \times F \times \rho \times \Delta \times E$
- All six canonical tokens: XP, CT, CAT, IT, DT, EP
- Validators, sensors, and the consensus layer that prevents forgery
- The decentralized ledger
- Rarity tables, domain weighting, decay schedules

Hard rules at Layer 3

- R is Rarity, not reputation. Past actions never inflate new XP. Fixed in v3.1.2.
- F is Frequency-of-decay. Not falsifiability. Not feedback closure strength. Not vote share.
- ρ (reputation density) lives only in CT. Reputation enters one place, not everywhere.
- Six tokens, no more. XP, CT, CAT, IT, DT, EP. No GT. No RT.
- Validators witness, they don't authorize. A user can't choose their validators. Validation is environmentally assigned.

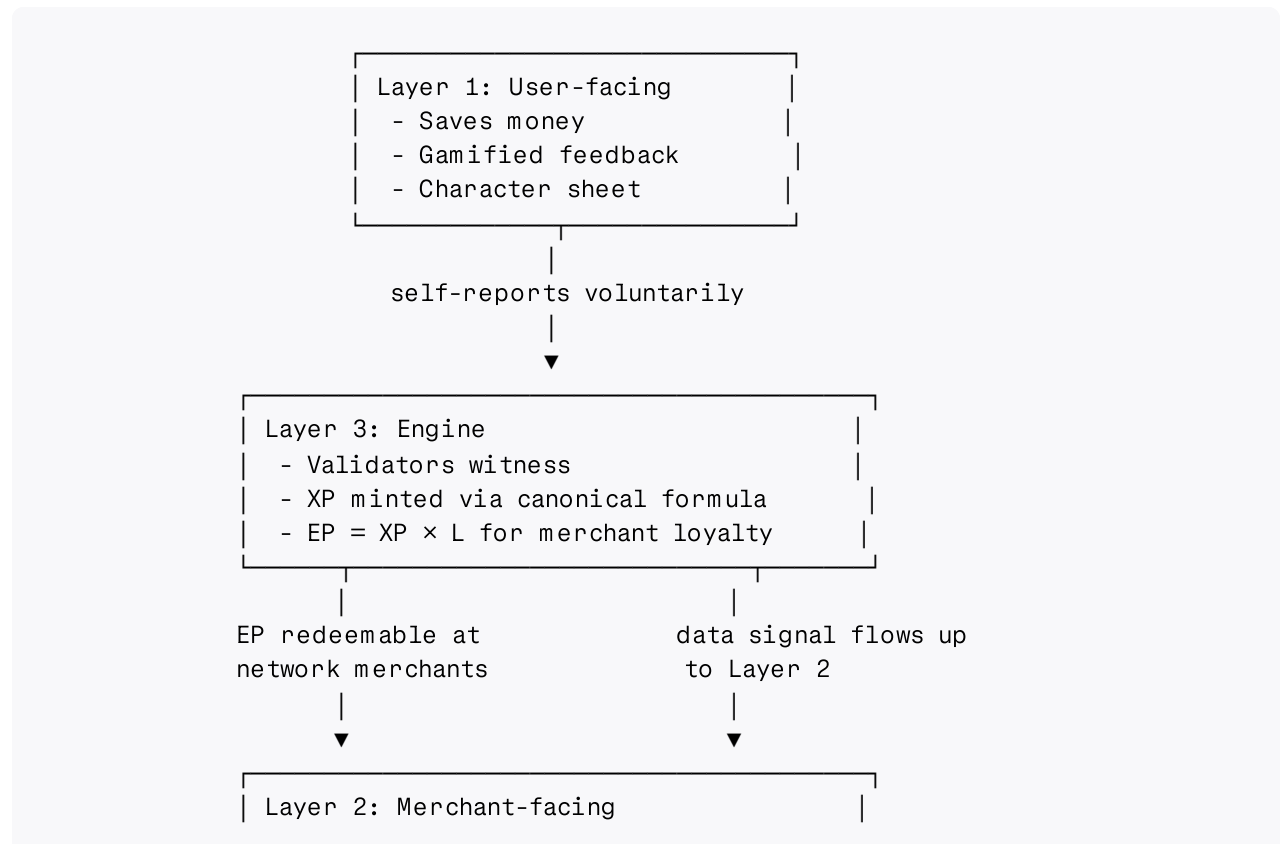
Why Layer 3 is not user-facing

Because the moment you show a user the formula, they will:

1. Try to maximize R by faking rare actions.
2. Try to maximize F by gaming decay schedules.
3. Try to maximize ΔS by inflating before-states.
4. Try to maximize $(w \cdot E)$ by claiming false domain coverage.
5. Try to minimize T_s by rushing.

Validators catch most of this. But the deeper defense is that users don't know which variable they're pushing. They see "complete this loop" or "you cooked at home tonight." They don't see the term in the equation that produces.

The flywheel



- Free POS
- Customer pipeline
- Operational signal
- Standard merchant fees captured

The user saves money. The merchant gets customers and better tools. The engine quietly does the actual coordination work. Nobody on either visible side has to understand the engine for the engine to function.

What this dodges

Goodhart's Law

The visible metric is not the rewarded metric. Users can't game what they can't see, and the proxy they *can* see doesn't pay out.

Surveillance capitalism

Self-reporting is voluntary. Users hold the keys to their own sheet. The state, an employer, or a corporation cannot write to a user's record. Only validated reality can.

Black Mirror social credit

The user holds the pen and the eraser. Reality holds the dice. There is no central authority that can downgrade a user. There are only validators witnessing what's already true.

Engagement-trap loyalty programs

The reward isn't engagement, it's actual money saved. A user who participates honestly once a month gets paid out for that. A user who farms an arbitrary engagement metric gets paid in dopamine, not dollars.

The privacy tax

The math is structured so that under-reporting honest signal does not penalize the user beyond the unrewarded action itself. R is per-domain rarity, ΔS is loop-specific entropy delta. Neither scales with disclosure volume. Sharing more does not, by itself, earn more. Doing more does.

Naming conventions

Internal (this doc, repo, dev branches, manifesto, the book)

- "Unfuck the world"
- "Unfucking [X]"
- Profanity is fine
- Voice is sharp, irreverent, technically literate, and unapologetic

External (POS software, merchant pitch, consumer marketing)

- "Rewire the everyday economy"
- "Repair the system from underneath"
- "Reward what actually matters"
- "Better incentives. Real savings."
- Profanity is out
- Voice is confident, dry, sharp, but won't get the project blocked from a school PTA Facebook page

The substance is identical. Two registers, one project. Don't confuse the audiences.

Hard invariants for any future contributor

1. Never expose raw XP to the user as a target number.
2. Never let a user choose their own validators.
3. Never let reputation re-enter XP. R is Rarity. Forever.
4. Never paywall the user-facing app.
5. Never charge merchants SaaS. Capture merchant-services fees instead.
6. Never mix the registers. Public copy stays public-safe. Internal copy can be as feral as it needs to be.
7. Never let self-disclosure volume scale reward magnitude. Privacy is not a tax.
8. Never let the gamification on Layer 1 become the actual scoring function. That's the trap.

If a feature, copy change, or contract change violates one of these, it doesn't ship.

Open questions

These are unresolved and worth tracking.

- Best front-end shape for the user? Browser extension, mobile app, web app, or a thin client across all three? Open.
- How aggressive should validator-side anomaly detection be? Tradeoff between false-positive friction and gaming resistance.

- What's the exact merchant fee structure at scale? 1.8% is a rough target. Real number depends on processor partnerships.
- Customer support staffing as a one-person project? Currently best-effort, volunteer. Document this honestly until funding allows otherwise.
- DFAO benefit ladder. What does a merchant get at each tier of node participation?

Closing

This document is the contract. The math, the tokens, the formula version, the merchant model, the user UX, the marketing voice — they all answer to this separation.

Three layers. Don't mix them. Don't merge them. Don't let anyone tell you the user "deserves to see their score." They don't. They deserve the savings. The score is the engine's problem.