Veridonia

Veridonia: A More Human-Centric Online Feed.

Hlib Semeniuk

Abstract

Veridonia is an experiment in building an online feed platform around a different objective than the one implicit in most existing feed designs. Editorial feeds answer what editors judge important; engagement-based ranking answers what will capture attention now; chronological feeds answer what was posted most recently. Each answers a coherent question, but none directly answers the question a platform would need to answer if its priority were to allocate visibility in a way that is representative of a community: **what each community as a whole decided should be seen by others today**.

This paper argues that answering that question requires treating feed curation as a governance problem under uncertainty: a system for allocating scarce collective attention through a procedure that is legible, contestable, and fast enough for daily use. Veridonia proposes a **referendum-like feed**—a claim about a process that approximates what a community would decide without requiring everyone to vote on everything. It combines randomised voter selection, majority voting, rating system, multi-stage voting proccess, and transparency and auditability into a pipeline for deciding what advances to visibility.

1. Introduction

Feed curation is a choice about how a platform allocates attention, and therefore about what kinds of contributions it rewards over time. When that choice is implicit or unstable, it becomes difficult to justify and contest outcomes: users cannot easily answer why something appeared, why something did not, or what behaviour the system is rewarding.

In practice, large platforms have converged on engagement as the dominant optimisation target. If revenue maximisation is the primary objective, this is a coherent design choice; the cost is that maximising engagement is not the same as maximising the value a user receives given limited time and attention. The result is familiar: sensational and emotionally charged material tends to outcompete slower, context-rich contributions, and misinformation can spread cheaply and quickly. (1, 2, 3) When ranking and moderation are opaque, it also becomes difficult to contest how visibility is allocated. (4)

Chronological feeds are sometimes proposed as a corrective because they avoid direct engagement optimisation. This is a partial improvement, but it does not solve the allocation problem under scarcity. Attention remains limited, and prioritisation is largely pushed onto the user. In high-volume communities, recency is a weak proxy for value. Editorial approaches address allocation more directly, but do so by centralising influence in a relatively fixed decision-making class.

Veridonia starts from a different premise. If a platform's priority is to serve its communities rather than to maximise engagement, the question should be different: what each community as a whole decided should be seen by others today. We refer to a system that tries to answer that question,

procedurally and at scale, as a referendum-like feed. Section 2 makes the contrast with existing feed designs explicit, and the rest of this paper describes the mechanisms required to implement it without presuming epistemic agreement or demanding universal participation.

2. Problem Statement

As a baseline, common feed designs can be described in terms of the questions they implicitly answer:

1. **Engagement-based feeds**: What do users pay the most attention to?
2. **Chronological feeds**: What was posted most recently?
3. **Editorial feeds**: What do editors deem important?

The problem is not that these are incoherent. It is that none is designed to answer the question a user is implicitly asking when they open a feed to be informed: what is most worth seeing given limited time and attention. In high-volume communities, this misalignment also makes the resulting allocation of visibility difficult to justify and difficult to contest.

The systemic consequences of these baseline designs are well-known:

Engagement-based feeds amplify what reliably captures attention, steering visibility toward emotionally charged, sensational, or outrage-inducing posts over slower, context-heavy material. (1, 2) Misinformation can spread faster than verification, and personalisation can concentrate users in ideological echo chambers. (1, 2, 3) When ranking and moderation are opaque, the resulting allocation of visibility is difficult to audit or contest. (4) These patterns correlate with downstream harms to well-being when feeds become saturated with high-intensity, low-signal content. (5)

Chronological feeds avoid direct engagement optimisation, but fail differently at feed scale. They offload prioritisation onto the user, turning the feed into a firehose in high-volume communities, and they produce path-dependent visibility where timing and early attention can bury high-quality contributions. In other words, "fair ordering" does not reliably produce "good allocation."

Editorial feeds provide a clear decision procedure, but centralise influence and scale poorly across diverse, fast-moving communities. They replace broad representation with a persistent decision-making class, which can be effective in some contexts but does not generalise as a community-governance model for everyday feed allocation.

Platforms are sometimes described as "democratic" because users can vote on content. But even minimal democratic systems have basic procedural attributes: (1) a defined question, (2) a defined electorate, (3) equal vote weight, and (4) a clear outcome. On Reddit, for example, the question an upvote is answering is often ambiguous (agreement, quality, relevance, humour, signalling), and there is little cost to treating those meanings as interchangeable. The electorate is also ill-defined in practice: outcomes are shaped by who happens to see a post early, who is online at the right time, and increasingly by automated participation. Vote weight is not equal, because early votes disproportionately shape visibility and moderation powers act as a persistent veto. Finally, the outcome is not a single decision with closure but a ranking that can be overridden by later attention or moderator intervention. The point is not that Reddit should be a referendum. It is that "voting exists" is not enough to make visibility allocation democratic in any minimal procedural sense.

Taken together, these dynamics describe a persistent allocation failure. Engagement-based ranking privileges what captures attention; chronological ordering pushes prioritisation onto users and makes outcomes sensitive to timing; editorial allocation can be coherent but concentrates influence

in a relatively fixed decision-making class. In each case, visibility becomes difficult to justify and difficult to contest, and high-signal contributions can be buried by volume, path dependence, or centralised judgement. Section 3 translates these failures into concrete design goals: the properties a feed should satisfy if it aims to allocate visibility more representatively and more robustly under feed constraints.

3. Design Goals: What a More Human-Centric Feed Should Be

A feed is not just a ranking function. It is a system for allocating scarce collective attention under uncertainty. Veridonia's technical design is meant to approximate the following properties; these define what "more human-centric" means in this paper and make the proposal testable. The mapping from these goals to Veridonia's five pillars is given in Appendix A.

3.1 Representative under uncertainty

A feed should reflect what a community would broadly consider worth attention without requiring everyone to participate in every decision. Representation should be statistical rather than exhaustive: under repeated random sampling, outcomes should be broadly stable and not depend on a small set of persistent decision-makers.

3.2 Resistant to capture, not dependent on trust

A feed should remain usable even when some participants act strategically, maliciously, or in coordination. Safety should come from structure rather than assumed good behaviour: influence should be costly to acquire and maintain, and attempted capture should leave detectable traces in voting patterns and rating movements rather than producing silent, stable control.

3.3 Open to participation, selective in influence

Anyone should be able to participate, but no one should have unlimited or permanent influence. A feed-scale system should separate voice from sustained weight: new users can participate immediately, while high-impact roles should churn over time and remain conditional on continued performance rather than early accumulation.

3.4 Self-correcting over time

A feed should be able to admit error, change its mind, and reallocate influence without manual intervention or regime change. Influence should be reversible: users (including editors) should be able to lose power through performance, and communities should be able to evolve norms without replacing governance.

3.5 Signal-prioritising by incentive, not intention

A feed should not depend on users being wise, informed, or altruistic. It should reward behaviours that produce signal regardless of motive, incentivising participants who seek influence to invest effort in judgements that generalise across the community rather than to maximise engagement or factional mobilisation.

3.6 Legible and contestable

Participants should be able to form a mental model of why content appears or does not appear. Decisions should be challengeable without exiting the system: users should be able to answer "why did I see this?" and "why didn't this appear?" with reference to concrete procedures, and appeals should be visible and tractable.

3.7 Fast enough for daily use

A feed that is correct but slow is functionally unusable. Speed is a constraint: for typical communities, time-to-decision should fit daily usage patterns, and per-user review load should stay bounded as content volume grows.

3.8 Procedural, not epistemic

A feed should not claim to know what is true, important, or correct. It should claim only that its process is fair, inspectable, and adaptive, so disagreements can be addressed by disputing procedure (sampling, rules, appeals) rather than by appeals to authority or opaque optimisation.

The sections that follow map these goals to a concrete mechanism set. Section 4 introduces Veridonia's five pillars and describes how they are used to implement a referendum-like feed under practical constraints.

4. Veridonia's Proposed Solution

Veridonia introduces a community-driven approach to structuring visibility in its online feed. The goal is to implement a referendum-like feed: a procedure intended to approximate what each community as a whole would decide should be seen by others, under constraints of scale and limited attention. The pillars below describe the mechanism set used to meet the design goals in Section 3—how participants are sampled, how decisions are made, how influence is earned and revoked over time, and how the process remains inspectable. Appendix A summarises how each design goal maps to one or more pillars.

1. Sortition (Randomized Participant Selection)
Random sampling broadens representation and limits coordinated control, ensuring that no fixed group consistently decides outcomes and that each decision reflects a changing cross-section of the community.

2. Consensus (Majority-Based Decision-Making)
Simple majority outcomes determine whether a piece of content advances, anchoring decisions in shared community standards rather than opaque algorithmic prediction.

3. Prediction-Based Rating System (PBRS)
A dynamic rating captures how reliably a participant's past decisions have matched community outcomes and turns that history into a notion of reputation. This reputation signal governs who is eligible for which responsibilities across the system—for example, participation in higher-impact reviews, moderation roles, and looser throttling.

4. Multi-Stage Voting Process (MSVP) for Post Publication
Tiered voting structures how posts move toward publication in community feeds. Early checks expose posts to a broad, low-cost sample, while later checks use smaller panels drawn from higher-rated

participants to approximate the outcome of a much larger community vote with far fewer total ballots.

5. Transparency and Auditability
Every moderation action, vote tally, and rating adjustment is publicly visible. This shifts trust away from assumptions about correctness and toward verifiable process, and allows communities to inspect how influence is earned and exercised.

5. System Architecture

The following components detail the implementation of Veridonia's five foundational pillars.

5.1 Submission & Review Pipeline

Veridonia evaluates each post that could appear in a community feed through a single, transparent pipeline that combines sortition (random sampling) and tiered majority voting. Random selection at each stage promotes fairness and diversity; tiering by rating concentrates final authority among proven reviewers without excluding broader participation. The mechanism scales with community size: as the population grows, random selection becomes harder to game; in very small communities the system collapses to a simpler single-stage vote.

The structure of this pipeline is chosen to balance three goals: keep decisions representative of the broader community, minimise the number of people who need to vote on any given post, and keep decision latency compatible with a live feed. The concrete stages below are a minimal arrangement that preserves diversity in early checks while concentrating later effort on a smaller set of participants who have demonstrated reliable judgement.

Process Flow

1. **Submit**: The author submits a post to a specific community.
2. **Stage 1: Initial Filter (lower 70% by rating)**: A random sample from the lower 70% by rating within that community reviews the post for relevance, informational value, and community alignment.
   Outcome:
   - If a simple majority approves, the post advances to Stage 2.
   - If a simple majority rejects, the post is rejected.
     Rating: After this decision, rating adjustments are applied to Stage-1 participants independently of Stage-2 outcomes.
3. **Stage 2: Final Decision (top 30% by rating)**: A random sample drawn from the top 30% by rating issues the final decision by simple majority.
   Outcome: The post either enters the community feed or does not, depending on the majority decision of the selected reviewers.
   Rating: Rating adjustments are applied to Stage-2 participants after the final decision.
4. **Small-Population Mode**: For communities with fewer than 20 members, a single random sample is drawn from all available users. A simple majority decides whether to publish. Rating adjustments are applied to those participants.

Rationale and Manipulation Resistance

- **Sortition:** Random selection reduces the viability of targeted manipulation and collusion.
- **Tiering by rating:** Final decisions are made by users who have consistently shown good judgement in filtering for relevance and quality within the community's scope, while keeping early checks broad to reflect community diversity and support scalability.
- **Efficiency under feed constraints:** For publication decisions about posts, Veridonia uses a two-stage pattern (Section 5.3) that approximates the decisions of a large one-stage community vote while keeping per-post latency and voter load compatible with a live feed.
- **Scalability:** Larger communities increase the entropy of selection, making coordinated capture more difficult.
- **Transparency:** All votes, outcomes, and subsequent rating changes are publicly logged for auditability.
- **Internal Echo Reduction:** Within a single community, the combination of random selection and majority outcomes tends to reward content that can attract support across factions. This pushes curation toward broadly acceptable signals and dampens the formation of narrow internal echo chambers, while still allowing distinct communities to maintain their own standards.

5.2 Prediction-Based Rating System (PBRS)

As one of Veridonia's five foundational pillars, the prediction-based rating system reflects how consistently a participant's prior decisions have matched the outcomes produced by their community. Over time, this identifies contributors whose votes are empirically predictive of full-community outcomes across many decision types. In practice, Veridonia implements this reputation score using an **ELO-style update rule**: after each decision, rating is transferred (zero-sum) from the losing side to the winning side, scaled by how "expected" the outcome was given each side's average rating. Higher-rated participants are invited into more consequential review stages, not as arbiters of correctness, but as members whose past participation suggests reliability in navigating the community's expectations:

- **Dynamic Influence:** Users' ratings reflect their track record of decisions relative to community outcomes. As these ratings rise, users become eligible for expanded responsibilities—such as participation in Stage-2 review for post publication decisions described in Section 5.3 or, where applicable, appointment as editors. These roles carry more weight in the curation pipeline that governs what appears in the community feed, are fully auditable, and remain conditional on continued performance.
- **Zero-Sum, Team-Weighted ELO Updates:** After the final decision, voters split into two teams: **winners** (their vote matches the outcome) and **losers** (their vote does not). Rating is reallocated zero-sum between these teams and weighted by their relative strength:
  1. Compute each team's average ELO rating (winners_avg, losers_avg).
  2. For each participant, compute an update scaled by a constant **K** and the gap between team averages. Members of the **winners** gain rating, moving upward toward the opposing team's average; members of the **losers** lose rating, moving downward toward the opposing team's average.
  3. The sum of all gains equals the sum of all losses (zero-sum conservation).
  4. This team-weighted update reinforces effective group-level filtering, amplifying the influence of participants who align consistently with community outcomes and reducing that of those who do not.

Conserving total rating makes influence a scarce resource that can only be reallocated from less

predictive to more predictive contributors, rather than inflated across the board. Weighting updates by team strength means that the size of each rating transfer depends on the gap between the average ratings of the winning and losing sides: when a lower-rated group wins against a higher-rated group, the adjustment is larger than when the higher-rated group wins as expected.

Because rating is updated after every decision and across both stages of review, the boundary between lower- and higher-impact roles is permeable. Participants who begin in Stage-1 review can, through a sustained record of alignment with community outcomes, move into Stage-2 and eventually into editorial (moderation) roles, while those whose decisions repeatedly diverge from outcomes will see their influence contract. This continual re-evaluation stands in contrast to rigid, once-appointed moderator classes common on other platforms and is intended to support a more bottom-up, renewable form of authority.

The numerical example below illustrates how these small, bounded adjustments operate in a single vote.

Example: One Voting Stage

Suppose five users have been selected to vote on whether a suggested post A should be published to a community X.

Their initial ratings are **800, 755, 821, 798,** and **804.**
Three vote **Yes** (users 1, 4, 5) and two **No** (2, 3). The majority outcome is **Yes**.

Step 1: Compute team averages

$$\mu_W = (800 + 798 + 804)/3 = 800.67$$

$$\mu_L = (755 + 821)/2 = 788.00$$

Step 2: Expected score for winners

$$E_W = \frac{1}{1 + 10^{(\mu_L - \mu_W)/400}} = \frac{1}{1 + 10^{(788.00 - 800.67)/400}} \approx 0.518$$

Step 3: Rating transfer

$$K = 32$$

$$\text{Total gain for winners} = K \times (1 - E_W) = 15.4$$

$$\text{Total loss for losers} = -15.4$$

Step 4: Distribution

Each winner gains $(+15.4/3 = +5.1)$
Each loser loses $(-15.4/2 = -7.7)$

After this round, the new ratings are approximately:
(805, 747, 813, 803, 809)

Although this example describes only one isolated voting stage, the same mechanism repeats continuously across many decisions and participants. Each round transfers small amounts of rating between participants whose votes align with the outcome and those that do not, and over time these micro-adjustments accumulate toward a stable equilibrium. In simulation over extended runs, the system self-organises into a characteristic distribution of ratings—most users cluster around the mean, with smaller groups at the extremes corresponding to more and less consistently predictive reviewers. The figure below shows this emergent pattern.

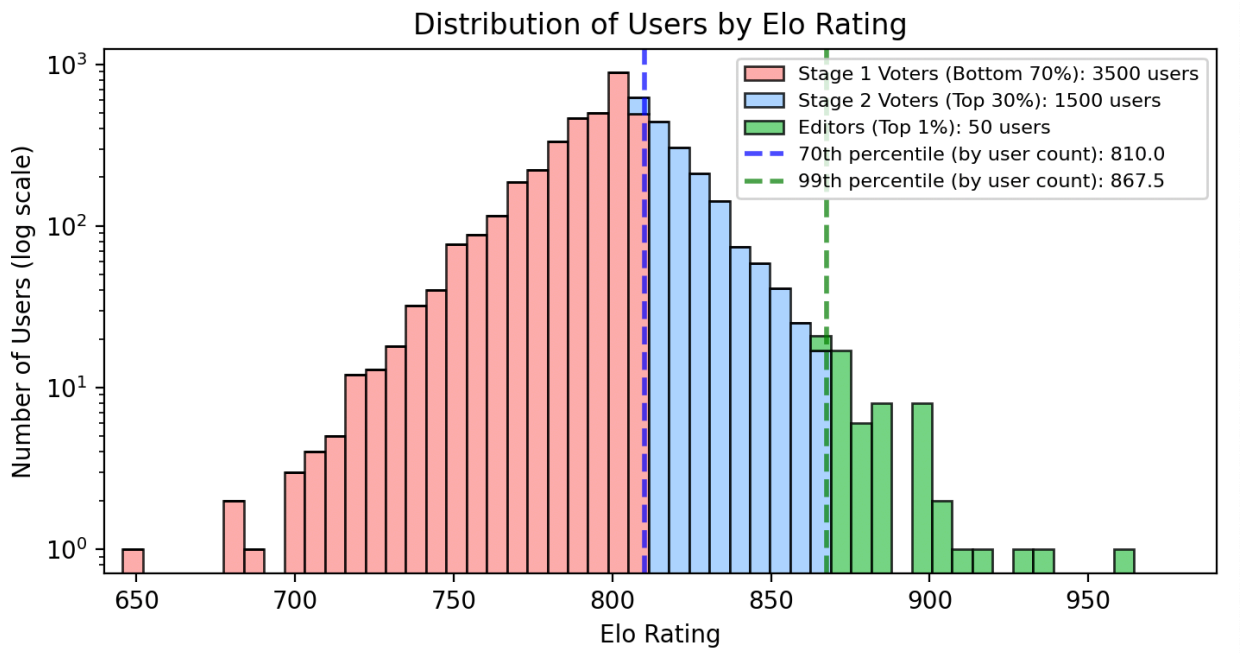

Figure 1: Distribution of Users by Rating. Simulation results showing the equilibrium state produced by repeated voting and rating updates. Most users cluster around the mean, with progressively smaller groups at higher and lower ratings corresponding to the initial and advanced decision groups. Dashed lines mark the 70th and 99th percentiles by user count.

Local Rating & Elected Cross-Community Stewards (TBD)
Rating remains **strictly local to each community**. A user's standing in one community neither boosts nor suppresses their standing in another, and no global rating is ever computed. This prevents the rise of "universal elites" and keeps influence contextual to demonstrated expertise.

In the future, Veridonia may introduce a small set of elected cross-community stewards ("chief editors") with limited administrative powers (e.g., emergency takedowns, cross-community maintenance). Details of their election, scope, and accountability are **to be determined** and may draw inspiration from Wikipedia's steward model. Crucially, these roles would not mix or merge community ratings, and routine content decisions would remain governed locally.

The rating system also encompasses onboarding and participation controls, detailed below:

5.2.1 User Onboarding and Baseline Attributes As a pragmatic defence against large-scale automated abuse, Veridonia currently couples initial user rating to an IP-level baseline while treating this mechanism as provisional rather than core to the system's philosophy.

- **Initial Assignment:** If the IP address has no previous users, a default rating (e.g., 800) is used.
- **IP-Based Rating Inheritance:** To protect the platform from bot attacks and coordinated manipulation, all new users inherit the rating assigned to their IP address. After each voting stage, an IP address is updated to reflect the lowest rating among its associated users. For example, if users from an IP (e.g., 156.156.156.3) have ratings of 850, 700, and 1500, the IP is assigned a rating of 700. Any new users registering from this IP will begin with a rating of 700, capped at a default maximum (e.g., 800) for first-time IPs.

While this IP-based inheritance mechanism mitigates certain manipulation risks, it has clear limitations. Shared or dynamic IP addresses may produce unintended effects—including the penalisation of legitimate users employing privacy-preserving tools (e.g., Tor or VPNs). This mechanism is not foundational to Veridonia's core philosophy; rather, it functions as an initial, pragmatic safeguard and is expected to evolve as the platform matures. Potential improvements under consideration include community-reviewed verification requests, whereby users could appeal or validate their onboarding status through review by top-rated participants (e.g., the top 30% or designated editors). The precise procedures and governance structures for such processes remain to be determined and will be shaped by community input and further research.

5.2.2 Rating-Based Throttling Mechanism Veridonia implements a throttling system that regulates users' ability to post based on their rating. Concretely, users with lower rating can post less frequently and experience longer cooldowns between contributions, and as their rating improves these limits are progressively relaxed.

The throttling mechanism serves several critical functions:

1. **Quality Control:** By limiting the volume of content from users with lower rating scores, the system naturally increases the average signal relative to noise in visible content.
2. **Spam Prevention:** Rate limiting creates an effective barrier against automated spam and coordinated manipulation attempts.
3. **Incentivizing Quality:** The direct relationship between contribution privileges and rating motivates users to focus on thoughtful, community-aligned contributions rather than quantity.
4. **Self-Regulation:** The system creates a natural self-regulatory environment where users who consistently provide low-quality content have diminished impact on the community.
5. **Resource Management:** Throttling helps manage computational resources by preventing system overload from excessive low-quality submissions.

This mechanism reinforces Veridonia's core principle that influence within the community should be earned through demonstrated alignment with community standards and quality contribution.

5.3 Multi-Stage Voting Process (MSVP) for Posts

Multi-stage voting is used specifically for publication decisions about posts that may enter community feeds. The central design question is how to approximate "what the whole community would decide" without asking a large share of the community to vote on every post.

As a reference point, one could imagine drawing a large random sample of the relevant community

for each post and taking a single majority vote. This would provide a direct snapshot of average opinion but would be prohibitively expensive for a live online feed: decision latency would grow, and participants would be overwhelmed by constant review demands.

The two-stage process is an optimisation of this baseline. Stage 1 uses a broad, randomly selected group drawn from the bulk of the community to filter out clearly off-scope or low-value submissions at low cost, preserving diversity and representation while reducing volume. Stage 2 then applies the same majority rule to a much smaller group of higher-rated reviewers whose ratings reflect a history of aligning with past community outcomes. Because these participants have repeatedly demonstrated that their judgements track what the broader community tends to decide, their votes serve as a sample-efficient proxy for a much larger community poll.

In expectation, this arrangement allows Veridonia to achieve outcomes that are comparable to, and on harder or more context-dependent posts potentially better than, those of a single large undifferentiated vote, while requiring far fewer total votes per decision and keeping latency compatible with an online feed. Other decision types—such as editors voting on maintenance proposals—may use a single stage, with each eligible participant carrying equal weight in that vote, while still relying on rating to determine eligibility and to update ratings after the fact.

6. Transparency and Self-Governance

Veridonia is designed to be an open and self-regulating ecosystem:

- **Public Auditability:** All voting records, rating adjustments, and moderation actions affecting what appears in the feed are logged and accessible for independent review, emulating blockchain-like transparency.
- **Decentralised Moderation:** Governance is vested in the community, with every member empowered to contribute, vote, and shape content standards for their feeds. Moderation rights are held by approximately the top 1% of users in a community by rating, who are able to soft-delete posts from the feed to uphold standards. Every moderation action can be appealed by the broader community through a randomized jury voting process.

Privacy by Design and Data Control:
Veridonia does not require sign-up to participate and does not track users across the web. The system only uses minimal signals necessary for fairness—for example, new accounts inherit the lowest rating from their IP to discourage bot farms. Beyond this, rating is tied entirely to actions within the platform: voting, posting, and how those decisions align with the community.

At the same time, users retain full control of their data. All activity histories can be accessed and verified without compromising security, reinforcing both transparency and individual privacy.

7. Additional Core Principles

Independence from Advertiser Influence: Veridonia is free from advertiser funding, ensuring that feed curation is not driven by advertiser incentives and is dictated solely by community standards. The platform will never employ advertising as a monetisation strategy. Instead, future revenue models may involve subscriptions or donations, but public benefit content—such as community feeds—will always remain free to access. Only private benefit features may be offered as paid options, balancing sustainability with Veridonia's commitment to open access and public good.

8. Conclusion

Veridonia is an experiment in community-guided online feeds. By pairing sortition and tiered voting with a prediction-based rating model (ELO-style), it replaces engagement optimisation with incentives that reward careful participation. Contributions that repeatedly fall outside the community's standards carry rating and throttling costs, while steady, attentive decisions expand a participant's role in shaping what the community sees in its feeds. The expectation is that such incentives produce feeds that feel more deliberate, legible, and aligned with the community's own preferences, with a healthier balance of signal over noise than engagement-driven alternatives.

The next step is empirical: testing the system under real conditions and deliberate stress. We will evaluate how feeds distribute attention between substantive contributions and noise, as well as overall content quality and capture resistance. We will also examine decision latency versus judgement alignment and the fairness of IP-based inheritance to refine both the model and its parameters.

To keep the claims in Section 3 falsifiable, we will measure (a) representativeness under uncertainty via outcome stability across repeated random samples and the concentration of decision power over time (Goals 3.1, 3.3), (b) capture resistance via the cost of influence acquisition, detectable coordination signatures in voting records, and the persistence of any captured state under continued participation (Goal 3.2), (c) self-correction via rating mobility, role churn, and reversibility of high-impact roles after shifts in community outcomes (Goal 3.4), (d) legibility/contestability via audit usage and appeal outcomes (Goal 3.6), and (e) daily usability via time-to-decision distributions and per-user review load under realistic volume (Goal 3.7). The procedural rather than epistemic framing constrains success criteria to these process-level properties rather than claims of objective correctness (Goal 3.8).

A second question is sustainability. We will evaluate whether a non-advertising model – driven by voluntary support or subscriptions – can fund operations without distorting incentives, while keeping core public-benefit features open.

Ultimately, Veridonia is a falsifiable proposal. If outcomes under real use do not beat practical baselines—or if funding compromises the aims—it should be revised or retired. If they do, the system may be worth iterating on. We invite researchers and communities to test, critique, and adapt these ideas.

9. Appendix A: Mapping Design Goals to Pillars

This appendix collects the mapping between the design goals in Section 3 and the five pillars introduced in Section 4.

Goal	Pillar	Contribution
Representation	Sortition	Samples a changing cross-section of the community rather than a fixed group
	Consensus	Aggregates sampled judgements into a clear, legible outcome
	MSVP	Approximates large-population outcomes with far fewer votes
Capture resistance	Sortition	Makes targeted influence harder by making reviewer selection unpredictable

Goal	Pillar	Contribution
	PBRS	Makes long-term influence expensive to maintain without continued alignment
	MSVP	Reduces attack surface by filtering volume early
	Transparency	Makes coordinated abuse more visible and investigable
Open participation	PBRS	Converts historical alignment into conditional, revocable influence and enables rating-based throttling
	MSVP	Routes final decisions to reviewers with demonstrated track records without excluding broader participation
Self-correction	PBRS	Reallocates influence after each decision, using zero-sum updates to prevent inflation and force trade-offs
	MSVP	Lets roles expand or contract dynamically as the rating distribution shifts
Signal incentives	Consensus	Keeps the optimisation target explicit and stable
	PBRS	Rewards accurate anticipation of community outcomes rather than declared intentions
	MSVP	Concentrates review capacity among participants who have repeatedly produced high-signal judgements
Legible & contestable	Consensus	Keeps outcomes contestable via reversible decisions (e.g., re-voting) under a clear rule
	Transparency	Makes decisions, vote tallies, and rating changes inspectable
Daily speed	Sortition	Avoids referenda-scale participation burdens
	PBRS	Reduces required sample size by concentrating decisions where they are most informative
	MSVP	Approximates large votes with fewer ballots and less delay
Procedural (not epistemic)	Consensus	Avoids hidden optimisation targets
	PBRS	Grounds influence in outcomes rather than credentials
	Transparency	Shifts legitimacy from assumed correctness to verifiable process

10. References

1. Pennycook, G., & Rand, D. G. (2021). "The Psychology of Fake News." *Trends in Cognitive Sciences*, 25(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007

2. Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., & Zittrain, J. L. (2018). "The Science of Fake News." *Science*, 359(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

3. Cinelli, M., Morales, G. D. F., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). "The Echo Chamber Effect on Social Media." *Proceedings of the National Academy of Sciences*, 118(9). https://doi.org/10.1073/pnas.2023301118

4. Ananny, M., & Crawford, K. (2018). "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability." *New Media & Society*, 20(3), 973–989. https://doi.org/10.1177/1461444816676645

5. Keles, B., McCrae, N., & Grealish, A. (2020). "A Systematic Review: The Influence of Social Media on Depression, Anxiety, and Psychological Distress in Adolescents." *International Journal of Adolescence and Youth*, 25(1), 79–93. https://doi.org/10.1080/02673843.2019.1590851