The Price of Forgery: A Gentle Methodology for Measuring Sybil Resistance

Petr Porobov
Upala Digital Identity

May 1, 2026

Abstract

There is no objective way to measure the quality of a Proof of Personhood (PoP) system. Designers and adopters compare methods through private trial-and-error or expert opinion. Yet many of the applications that need PoP the most — quadratic funding, DAO voting, basic income distribution, democratic governance of AI systems — would benefit from a shared yardstick to choose between methods or to combine them.

We argue that the only available objective measure is one set by the market, and propose the **price of forgery** (PoF) — the dollar amount required to forge an identity that a given Human Verification (HV) method would accept — as a natural form for that measure to take. PoF is objective because it is not anyone's opinion. It is set by the people who would actually do the forging.

This paper introduces the Upala protocol — an incentive system where a user can voluntarily mark their own identity as Sybil in exchange for money — and the Gentle Methodology, a cost-efficient measurement campaign that uses Upala to discover the PoF of any HV method. Once PoF can be measured, organizations can quality-control the methods they rely on, combine multiple methods into aggregated identities of higher value, and let an open market push HV methods to compete on price and security.

1 Introduction

1.1 Proof of uniqueness

Many systems work better when each account is one and only one person. Ad networks want one impression per human. Voting wants one ballot per citizen. Quadratic funding wants one donor per donation. None of these need to know *who* a user is — only that no one else is them. The problem of separating unique humans from copies of the same human is called **proof of uniqueness**, more commonly **Proof of Personhood** (PoP). A PoP method takes a user and returns a verdict on whether the user is a previously-unseen person, with some level of confidence.

Different applications need different levels of confidence. Showing an ad to a CAPTCHA-verified user is fine. Voting on a billion-dollar treasury is not. The cost of acquiring an identity should match the value that identity unlocks: cheap for ads, expensive for financial tools, somewhere in between for everything else.

1.2 State of the art in PoP is subjective

A new generation of PoP protocols has emerged over the past decade, each one a careful piece of mechanism design built on a different signal. BrightID derives personhood from vouched connections in a social graph and has been live for years, showing that a working Web of Trust can be assembled without biometrics, government IDs, or a central operator. Proof of Humanity pairs a token-curated registry with video submission and Kleros-style on-chain arbitration to adjudicate identity claims at scale [10]. The Idena Network runs synchronous, human-generated reverse Turing tests in a fully decentralized validation ceremony — no central authority anywhere in the loop, with all participants solving "flips" together at a fixed moment in time. PoPCoin [1] introduced the term "proof of personhood" itself and proposed a cryptocurrency whose consensus is anchored by pseudonym parties — periodic in-person events first sketched by Ford and Strauss [4], where uniqueness is established by a constraint the physical world enforces for free: a person can only be in one place at a time. Duniter also runs a Web of Trust, tied directly to a universal dividend currency. Earlier protocols from the Web 2.0 era — SybilLimit [12], SumUp [11], Whānau [5], and others — laid the theoretical groundwork, proving formal bounds on how many Sybils a social graph can absorb before collapsing [6]. Those results still underpin most of the social-graph designs that came after.

What none of these methods set out to provide — and what the field as a whole still lacks — is an objective, public measure of how secure each one is. Academic protocols were necessarily tested against synthetic graphs and simulated adversaries; deployed systems learn the real numbers privately, through incidents and patches that rarely surface outside the team. There is no shared figure that says *"this method costs $X to break"*. Each project that depends on Sybil resistance is left to estimate that cost on its own, using its own assumptions and its own incidents.

1.3 Why subjective is not enough

The real-world property that actually matters is the amount of value a single account can extract. CAPTCHA is enough for

ads. SMS verification is enough to gate a frontier AI model. A passport scan is needed for regulated finance. Behind each of these choices is somebody who priced the interaction and picked a verification method they thought could defend it. They got there through analysis and trial-and-error — and the result stayed private inside the company.

This is workable as long as nobody else needs to use the same numbers. Once an application has to combine methods, compare providers, or expose security claims to outside parties, private estimates stop being enough. Distributing universal basic income, running elections on-chain, governing AI systems democratically, allocating public goods funding — all of these need a *public* quantitative measure of how much an identity is worth.

1.4 Price of forgery

For these applications, it does not matter whether an account is a sybil, a puppet, a bot, or a real human having a bad day [3]. The only operational question is:

How much does it cost to forge an identity like this?

We call the answer the **price of forgery** (PoF). PoF is objective because it is not anybody's opinion. It is the price at which forging an identity becomes profitable for the cheapest attacker around. That price is set by the people who would actually do the forging, and they do not need to be polled. If the price is wrong, the market will correct it — somebody will take the money. The market cannot be argued with [7].

What is missing is a way to ask the market its price openly. This paper proposes one. We describe the Upala protocol — where any participant can voluntarily mark their own identity as Sybil in exchange for money — and a cost-efficient measurement campaign built on top of it. Together they turn *"is this user a real person?"* into a dollar number that any organization can measure and any DApp can read [2].

2 The Upala Protocol: An Incentives Model for Identity

Instead of preventing Sybil attacks by design, Upala provides a framework to empirically *measure* the economic cost of a successful Sybil attack against *any* existing Human Verification method.

This paper introduces the Upala protocol and a methodology designed to answer this question directly. Upala is an incentive-driven system that assigns users a score valued in dollars, representing an amount of money they can claim by destroying their identity. This mechanism opens a market for identity. The "Price of Forgery" (PoF) emerges from this market as a clear, dollar-denominated metric representing the security of a given Human Verification (HV) method.

Another contribution is a cost-efficient auction mechanism — the "Gentle Methodology" — for discovering the PoF. The campaign starts with small payouts and raises them gradually; when a coordinated attack appears, the bid that triggered it is recorded as the PoF of the underlying HV method.

2.1 How the protocol works

The building block of Upala is the **group**, a collection of users governed by a distinct set of rules — typically a Human Verification method that determines who is admitted. Groups assign **scores** to their members, which are denominated in dollars. A user's score is not merely a reputation metric; it represents a claim on real funds.

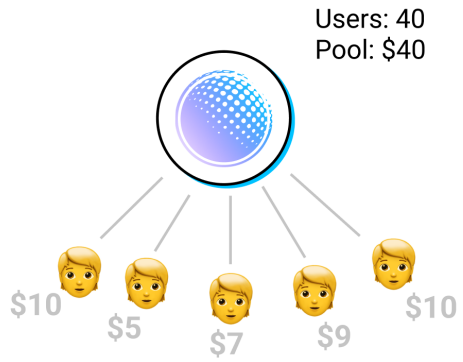

Figure 1: Groups assign dollar-valued scores to their members.

Each group maintains a **pool** of money. Any member of the group can, at any time, choose to liquidate their identity. This action is irreversible and allows the user to withdraw an amount of money from the group's pool equal to their score. The mechanism creates a powerful incentive for groups to assign scores carefully and to curate the HV method that gates admission, as the group's own funds are at stake.

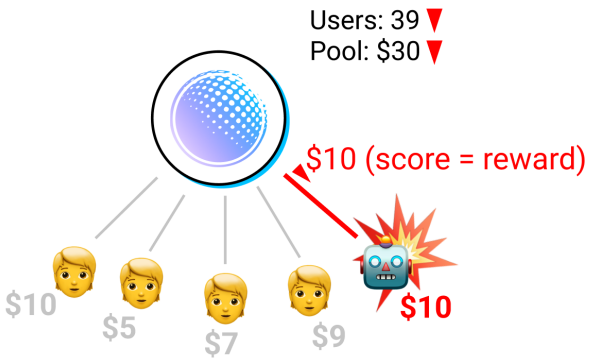

Figure 2: A user's score is the reward for self-liquidation.

The same setup looks much simpler from the bot farmer's side. A farmer evaluating a group has one question: does the score exceed the cost of forging the entry test? If yes, every fake identity is profit waiting to be liquidated. If no, the operation runs at a loss. Nothing else weighs in. The score at which

the answer flips is the price of forgery. Upala does not compute that number anywhere. It walks the score upward in small steps until the bot economy reacts, and records the bid at which the reaction happened. Everything else in the methodology is how to do that walk without spending more than necessary.

Even without a competitive landscape of multiple groups, this core mechanism is inherently useful. A single organization can create a group to quantitatively assess the security of its own internal HV methods. By funding a pool and running the PoF discovery methodology described later, the organization can determine its vulnerability to Sybil attacks in precise economic terms. This standalone utility for quality control is one of the primary applications of the protocol (see Section 5, Applications and Use Cases).

3 A Gentle Methodology for PoF Discovery

The PoF for an HV method can be measured empirically. The following methodology is designed to be cost-efficient, balancing the trade-off between the time taken and the capital risked during the discovery process. It is composed of two interconnected auctions.

3.1 The Score Auction

The first phase of the methodology is a "Score Auction," where we systematically determine the liquidation value that attracts bot armies.

Auction Overview

- **Score**: The amount of money offered to a user who has passed an HV method.
- **PoF**: The score value that we consider the current Price of Forgery. This is the unknown variable we are measuring.
- **Bid**: The specific Score being tested at any given time.

The measurement proceeds in discrete steps. The bid is set to a low initial value and is raised slowly over time. As soon as the number of bot liquidations within a period exceeds a predefined threshold, the current bid is considered to have revealed the PoF. The bid is then reset to a much lower value, and the gradual increase begins again to continuously track changes in the PoF.

Score Auction Parameters

- **Bid Duration**: The period of time during which a single bid is tested (e.g., one day).
- **Starting Bid**: The initial low score at which the auction begins.
- **Bid Step**: The percentage increase applied to the bid if the previous period did not meet the liquidation threshold. Smaller steps yield higher accuracy but require more time.

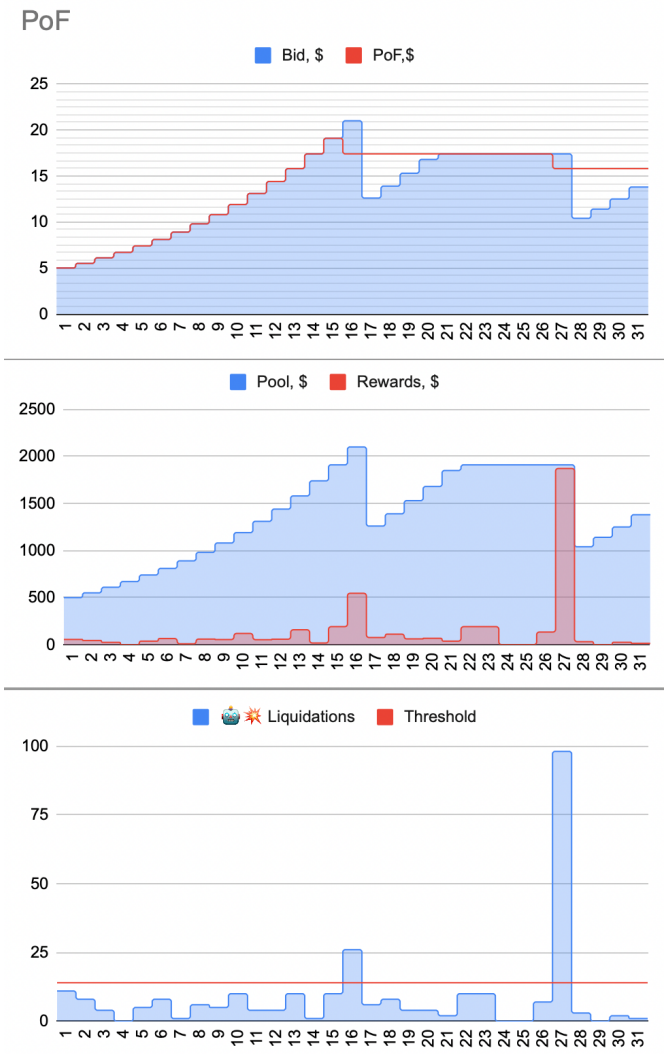

Figure 3: The Score Auction process, where the bid increases until the liquidation threshold is met, revealing the PoF.

- **Liquidations Threshold**: The number of liquidations (or total dollar value) that signifies a coordinated attack by a bot army, as opposed to random, non-malicious liquidations.
- **Step Down**: The percentage by which the bid is reduced after the threshold has been hit.
- **Pool Size**: The total amount of money at risk, which determines the number of bots (n) that can be liquidated simultaneously at a given bid:

$$n = \frac{\text{Pool Size}}{\text{Bid}} \tag{1}$$

3.2 Relationship Between PoF and Pool Size

The PoF is not a static value; it is intrinsically linked to the potential profit an attacker can realize, which in turn depends on the total capital available for liquidation (the Pool Size). The

financial model for an attacker can be expressed as follows:

$$\text{PoF} = \text{CoF} + \frac{(\text{Fixed\_Costs} + \text{Revenue})}{n} \tag{2}$$

where:

- **CoF** is the marginal Cost of Forgery per bot (e.g., labor, fees).
- **Fixed_Costs** are the attacker's initial setup costs.
- **Revenue** is the total profit for the bot farmer.
- **n** is the size of the bot army that can be liquidated.

An attacker's revenue is therefore:

$$\text{Revenue} = n \times (\text{PoF} - \text{CoF}) - \text{Fixed\_Costs} \tag{3}$$

This shows that an attacker's incentive is directly proportional to n, the number of bots they can liquidate. Since n is controlled by the Pool Size, the pool size itself becomes a critical variable in the discovery process. A pool of $10 is unlikely to incentivize any serious attack, whereas a pool of $1,000,000 might incentivize the development of novel exploits.

3.3 The Pool Size Auction

To account for this relationship, we introduce a "Pool Size Auction" that runs concurrently within each bid duration of the Score Auction.

Instead of keeping the Pool Size static, we gradually increase it during a single bid period. For example, over one day, we might increase the pool from $1,000 to $5,000 in discrete steps.

If the liquidation threshold is met, we record not only the PoF (the current bid) but also the Pool Size at which the attack occurred. This provides a much richer data point: the precise capital risk required to trigger an attack at a specific price point. If the threshold is not met by the time the pool reaches its upper limit, we advance to the next bid in the Score Auction and reset the pool to its lower limit.

Pool Size Auction Parameters

- **Lower/Upper Pool Size Limit**: The starting and maximum pool size for a given HV method test.
- **Number of Steps**: The number of times the pool size is increased within a single bid duration.
- **Pool Size Step**: The value of each increase, which can be calculated as:

$$\text{Pool\_Size\_Step} = \frac{(\text{Upper Limit} - \text{Lower Limit})}{\text{Number of Steps}} \tag{4}$$

Careful selection of these parameters is crucial for balancing the accuracy of the measurement against the cost of the campaign.

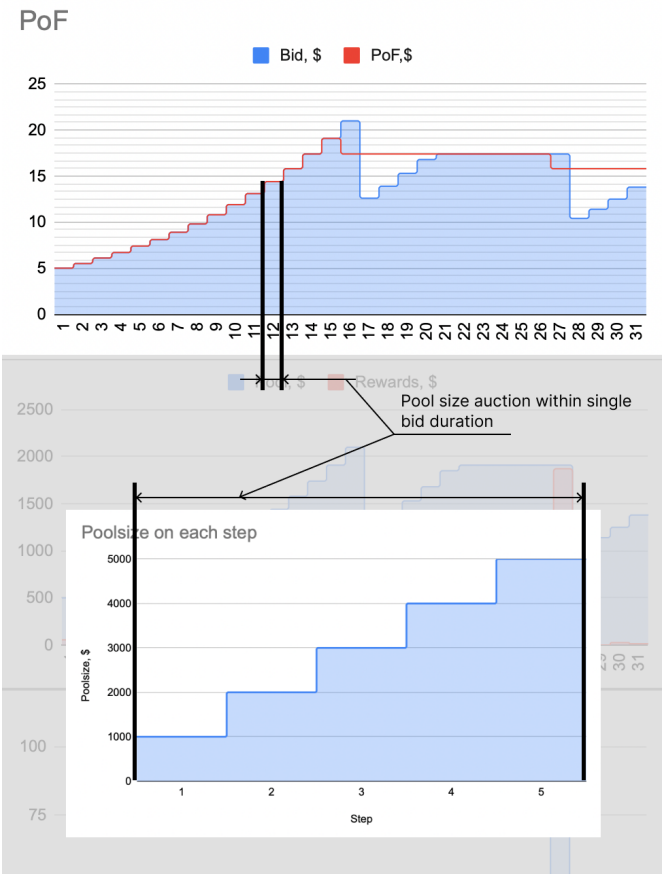

Figure 4: The Pool Size is increased in steps during a single bid duration, with liquidations potentially triggered at a specific size.

3.4 The Emergence of a Free Market for HV Methods

The mechanics described so far are the whole protocol. What follows is not. It is one economic arrangement the protocol makes possible — groups sell scores to DApps, DApps pay fees, and a market for HV methods emerges from the relationship. The protocol does not require it. The rest of this section assumes the conditions under which such a market takes shape: multiple groups running HV methods, DApps that need them, and a willingness on both sides to transact in scores.

When multiple groups and DApps adopt the protocol, a dynamic, free market for identity verification emerges. Decentralized applications (DApps) can leverage this system by choosing which groups to trust as score providers. To grant a user access or benefits, a DApp can verify the user's score, with the assurance that this score is backed by a tangible economic stake in the group's pool.

This model establishes a market where groups compete to attract users by offering the highest possible scores for the least amount of entry effort (e.g., passing an HV test). This competition benefits DApps by driving down fees and increasing the quality of available scores. However, this upward pressure on scores is counterbalanced by the constant threat of bot farmers.

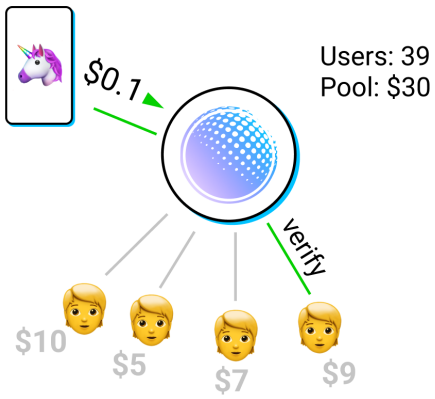

Figure 5: DApps nominate groups as their trusted score providers.

Bot farmers seek to exploit HV methods and liquidate identities for profit. They will attack any group whose score surpasses their cost of forgery. This equilibrium point, where the score is just below the profitability threshold for an attacker, is the Price of Forgery (PoF).

Two consequences follow. The PoF in any group is not a number set once and forgotten — it is recalculated continuously by the same forces. As HV methods harden, the equilibrium climbs; as new exploits surface, it falls. The number tracks the real security of the underlying method in something close to real time. And the same dynamics open a market for HV methods themselves: groups that can sustain higher scores at lower entry cost win more DApps and more users, which means the most effective verification methods accumulate revenue and the weakest are priced out. Sybil resistance becomes something that can be sold.

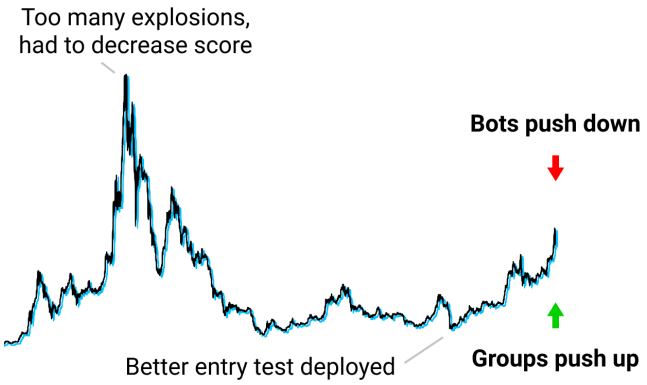

Figure 6: The score gravitates towards the Price of Forgery due to competing market forces.

4 On Common Misconceptions

The protocol has been described informally in talks and posts long before this paper, and a small set of objections recurs in conversations. We address them here.

4.1 "But you are paying bots."

The first reaction to Upala is almost always that paying money to bots is irrational.

But that dollar is the price of information. A pool drained at bid $X has answered a question that nothing else answers directly: *the human verification method protecting this group costs at most $X to forge.* Without that number, a project designing for Sybil resistance is guessing — overspending on defenses for low-value interactions, or underspending and discovering it only after a successful attack. The capital lost in a measurement campaign is bounded and predictable. The capital lost to a miscalibrated HV stack is neither.

A second framing follows from the first: in many configurations, paying out a measured PoF is *cheaper than fighting Sybils directly.* Detection systems are open-ended work — an attacker only has to win once, and HV methods that worked yesterday rarely work next year. A measured PoF lets a project decide what it can tolerate. If the value at stake in a single account is below the PoF, the attack is unprofitable for the bot farmer and no further defense is needed. If the value is above the PoF, the project has a number to size additional security against. In the cleanest case, the most cost-efficient strategy is not building a higher gate but making sure the prize on the other side is smaller than the price of forgery — bribing the bot economy away from the gate rather than reinforcing it.

4.2 The Bot Tragedy of the Commons

A related objection: by making forgery profitable, do we not encourage more of it?

The opposite turns out to be the case. A bot farmer participating in a PoF campaign is selling, in effect, the cost structure of their own operation. Every successful liquidation is a published data point that anchors expectations and informs the next generation of HV design. The same farmers who profit from one campaign contribute the data that makes the next HV method harder for them to forge.

This is closer to a bug bounty than to a black market. A vulnerability researcher who sells an exploit to a vendor is paid for raising the cost of future exploits. A bot farmer who liquidates in an Upala campaign is paid for raising the cost of forging the HV method they just defeated. The mechanism is adversarial in form but cooperative in effect: an open market for information that participants on every side benefit from.

4.3 Upala is not a Proof-of-Personhood method

A persistent confusion is that Upala competes with BrightID, Worldcoin, Idena, or similar. It does not. Upala does not verify users, run challenges, maintain a registry of unique humans, or hold an opinion about what a human is.

What Upala does is *measure the price* at which an existing PoP method's verdict can be bought. The HV method itself can be anything: government ID checks, biometrics, social graphs,

CAPTCHAs, video calls, retinal scans, or methods that have not yet been invented. Upala wraps each of them as a group, denominates their security in dollars, and lets a market discover the number.

That the resulting score *can* be used as a personhood signal — a DApp may decide to trust any user with a PoF above $100, for example — is a useful by-product, not the design goal. The design goal is measurement. Personhood-as-output emerges because once an identity has a public dollar score, downstream applications can do whatever they want with it, including treating high-score identities as personhood claims.

The distinction matters in practice. Upala is not in competition with PoP protocols; it sits underneath them. A Worldcoin orb verification, an Idena flip session, and a notarized passport scan can all be groups in Upala. A user holding three independent verifications has a PoF roughly equal to the sum of the three, since an attacker would have to forge all three to drain the user's stake. The standardized dollar denomination is what makes that combination possible at all, and it is what allows hybrid identities to reach a price of forgery that no single HV method could deliver on its own.

4.4 Will money rule? On oligarchy

A more philosophical objection: any system in which identity has a price is a system in which the rich can buy as many identities as they want. Does Upala bake oligarchy into the protocol? Siddarth et al. [10] raised this concern about Upala directly in their review of PoP protocols: scores denominated in dollars and backed by pools may favor capital-rich users.

Two responses. First, the bribery being measured is not bribery introduced by Upala. Every PoP method already has a price; the absence of an open auction does not mean the auction does not exist, only that it is happening privately, in markets the attackers run and the defenders cannot see. Solved CAPTCHAs are sold by the thousand. SMS verification rings exist. Government IDs have been forged for as long as governments have issued them. Upala does not create a market for forgery — it makes an existing one legible.

Second, legibility is itself a defense. A system whose forgery price is *unknown* can be quietly attacked by anyone with the budget to pay. A system whose forgery price is published can be defended by anyone with the budget to monitor it. The asymmetry that favors well-resourced attackers shrinks when the attack is no longer hidden. It does not vanish — capital still talks — but the conversation takes place in public, where defenders can hear it.

Upala does not promise to remove the influence of money on identity. It assumes that money already influences identity, in every system, and proposes to measure it.

4.5 Sybils, collusion, and one measurement

Sybil resistance and collusion resistance are usually treated as separate problems. Sybil is the personhood problem: prove this account is a unique human. Collusion is a mechanism-design problem: design rules so that unique humans cannot profit by working together. Different literatures, different defenses.

Upala measures both with the same auction, as long as the defense can be expressed as a well-defined flow. An HV method fits naturally: the user passes a check — credentials, a CAPTCHA, a biometric, a video call, anything — and the method admits or rejects them. An anti-collusion method fits whenever it has the same shape: a defined input, a defined rule, a defined outcome. Wrap that flow as a group's entry condition, fund a pool, and run the auction. The mechanics are identical to a Sybil-resistance campaign.

This means PoF measures something stronger than its name suggests. It is the cost of accumulating enough adversarial influence to extract value, *by any means* — forging accounts, bribing real users, hiring mechanical turks, coordinating a friend group. One number prices them all.

The protocol also sidesteps a debate that has consumed considerable energy in the personhood literature — where exactly Sybil ends and collusion begins. Idena's puppeteer crisis [9] was exactly this kind of edge case: a single operator paying validated humans for the use of their accounts blurs the line between one attacker with many identities and many users colluding. Upala does not need to take a position. It measures the price.

4.6 Avalanche exits

A specific structural concern, raised by Siddarth et al. [10] in their review of PoP protocols, is that Upala may be vulnerable to an *avalanche user exit*: if confidence in the protocol falters for any reason — a hack, a publicized attack, a rumor — a wave of users could race to liquidate before the pools empty, draining groups en masse. The shape of the worry is familiar from banking: a bank run.

The analogy breaks at the layer that matters. A bank holds a single pool of fungible deposits and a coordination problem across all of its depositors at once. Upala holds many independent pools, one per group. Loss of confidence in one group does not drain another. There is no shared reserve to run on, no protocol-level token whose collapse cascades. If confidence in a particular HV method evaporates, the worst case is that its pool empties and that group's PoF resets to whatever the market thinks it is now. That is not a failure mode — it is a measurement.

Two further asymmetries weaken the run within any single group. Liquidation is destructive: a user loses their identity in that group permanently and forfeits any future score, so the threshold for joining a panic is meaningfully higher than for withdrawing from a bank. And a user's score is a claim against the pool only — if the pool empties, latecomers get nothing and *keep* their identity, which still has value in any other group whose pool is still funded. The supply of liquidators is bounded by who gets there first; everyone else goes home with the asset they came in with.

The upper bound on the worst case in any single group is the size of its pool — a number group operators choose deliberately when they fund it. The word "avalanche" suggests

something that grows beyond its initial mass; what Upala actually exposes is the exact mass each group has elected to put at risk.

5 Applications and Use Cases

The ability to assign a clear, dollar-denominated security score to an identity system unlocks numerous applications. Some are available today with the protocol as it stands. Others come into reach if multiple HV methods adopt the protocol, and even more when a competing market emerges.

5.1 Available today

- **Quality Control for HV Methods**: Systems that depend on Sybil resistance, such as Gitcoin for quadratic funding [8] or DAOs for voting, can use PoF to compare different HV providers and select the one that offers the best security for the cost.
- **Dynamic Security Thresholds**: A DApp can set a minimum PoF score for users to access certain features. A user's score becomes a quantifiable measure of how much a trusted group is willing to stake on their authenticity.
- **Identity as Collateral**: Lending platforms could accept a user's Upala ID as a form of collateral. If the user defaults, the platform can liquidate the ID to recoup its losses.
- **Testing Assumptions**: A team building a new HV method can wrap it in an Upala group and run a small PoF auction before launch. The auction settles on the price at which someone is willing to forge an identity through the method. If that price comes out lower than what the team assumed, the economics fail in a controlled measurement rather than in production. Idena could have wrapped its validation ceremony this way and seen its puppeteer crisis priced in advance: by 2022, a single operator could obtain a validated Idena account for roughly $20 a month, and a small handful of such operators came to control around 40% of the network — a price an Upala auction would have surfaced long before it became visible in the open [9].

5.2 As the market grows

- **Monetizing Identity Systems**: Any existing identity system, from a government ID provider to a social network, can be wrapped in the Upala protocol and monetize its verification services by providing PoF-rated scores to DApps — the score-provider role described in Section 3. Open competition between providers creates pressure for stronger HV methods, exactly as in the market dynamics described above.
- **Cross-Chain Identity Bridge**: Upala can be deployed on any blockchain. A score provider native to one chain can register as a group on another, so DApps on the inviting chain read its identities through a local group exactly as they would for any HV method that originated there. The protocol becomes a common denominator for identity across ecosystems, regardless of where each HV method lives.
- **Identity Aggregator**: Since Upala is a standard, multiple HV methods can be combined into a single ID, with no fixed ceiling on the resulting PoF. Combining a government ID verification with a PoPCoin verification, for example, produces an identity worth more than either method alone.

6 Conclusion

6.1 A field without a number

Proof of Personhood has run for over a decade without a way to measure itself. The applications that depend on Sybil resistance — quadratic funding, DAO treasuries, basic income, governance over AI — work with budgets in the millions and pick HV methods on intuition.

6.2 Price of forgery

The argument of this paper is that an objective measure of Sybil resistance has to be set by the market — no committee or expert panel will produce one — and that the price of forgery is a natural form for such a measure to take. It is the dollar amount required to produce an identity that a given HV method will accept, set by whoever is cheapest: bot farmer, document forger, mechanical turk operator, social engineer. Whichever of them is willing to take the money sets the number. We offer PoF to the Sybil-resistance literature as a candidate yardstick — dollar-denominated, comparable across methods, set by the people who would actually attack.

6.3 A methodology to measure it

The paper also describes a way to discover the price. The Upala protocol — an incentive layer in which any user can liquidate their own identity for the score a group has staked on them — together with the Gentle Methodology, a dual auction over score and pool size, walks the market upward in small steps until forgers reveal the number. The campaign's cost is bounded and known in advance, and for applications whose stakes exceed the current PoF, the math tends to favor measuring. Once a number exists, thresholds can be set in dollars instead of trust, and HV methods can be compared and combined on common terms.

Thank you for reading, human.

References

[1] Borge, M., Kokoris-Kogias, E., Jovanovic, P., Gasser, L., Gailly, N., & Ford, B. (2017). "Proof-of-Personhood: Redemocratizing Permissionless Cryptocurrencies." In *2017 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, 23–26.

[2] De Filippi, P. (2019). *Blockchain Technology and Decentralized Governance: The Pitfalls of a Trustless Dream.* SSRN Scholarly Paper ID 3524352. Rochester, NY: Social Science Research Network.

[3] Douceur, J. R. (2002). The Sybil Attack. In *Revised Papers from the First International Workshop on Peer-to-Peer Systems* (pp. 251–260). Berlin, Heidelberg: Springer-Verlag.

[4] Ford, B., & Strauss, J. (2008). "An Offline Foundation for Online Accountable Pseudonyms." In *Proceedings of the 1st Workshop on Social Network Systems*, 31–36. SocialNets '08. New York, NY: ACM.

[5] Lesniewski-Laas, C., & Kaashoek, M. F. (2010). "Whānau: A Sybil-Proof Distributed Hash Table." In *Proceedings of the 7th USENIX Symposium on Networked Systems Design and Implementation*. NSDI '10.

[6] Margolin, N. B., & Levine, B. N. (2008). "Quantifying Resistance to the Sybil Attack." In *Financial Cryptography and Data Security*, LNCS vol. 5143. Berlin, Heidelberg: Springer-Verlag.

[7] Mazorra Roig, B., & Della Penna, N. (2023). *The Cost of Sybils, Credible Commitments, and False-Name Proof Mechanisms.* arXiv preprint arXiv:2301.12813.

[8] Miller, J., Weyl, E. G., & Kanich, C. (2025). "Fair Decisions through Plurality: Results from a Crowdfunding Platform." *arXiv preprint arXiv:2509.18343*.

[9] Ohlhaver, P., Nikulin, M., & Berman, P. (2024). *Compressed to 0: The Silent Strings of Proof of Personhood.* SSRN Scholarly Paper ID 4749892.

[10] Siddarth, D., Ivliev, S., Siri, S., & Berman, P. (2020). *Who Watches the Watchmen? A Review of Subjective Approaches for Sybil-resistance in Proof of Personhood Protocols. Frontiers in Blockchain*, vol. 3.

[11] Tran, D. N., Min, B., Li, J., & Subramanian, L. (2009). "Sybil-Resilient Online Content Voting." In *Proceedings of the 6th USENIX Symposium on Networked Systems Design and Implementation*, 15–28. NSDI '09.

[12] Yu, H., Gibbons, P. B., Kaminsky, M., & Xiao, F. (2008). "SybilLimit: A Near-Optimal Social Network Defense against Sybil Attacks." In *2008 IEEE Symposium on Security and Privacy*, 3–17.