Nown: A Peer-to-Peer Protocol for Portable Trust and Merit-Based Arbitration

Anonymous
nown.to github.com/nownto/nown

Draft v0.3

Abstract: Trade between strangers requires a way to establish trust and a way to resolve disputes. Every market solves both the same way, with a trusted third party: a platform, an escrow agent, a rating bureau, or a court. The third party can be censored, bribed, subpoenaed, or shut down, and it owns the reputations and the verdicts it issues. This protocol removes the third party from both functions. Reputation, called karma, is a portable measure each person earns through costly action and carries between contexts. It proves baseline trustworthiness to any stranger, with no introduction, no identity document, and no platform. Disputes are decided by a quorum of high-karma accounts drawn at random, anonymous to the traders and to each other, judging only the evidence each side submits. The quorum is an oracle. It attests a verdict that a pre-signed escrow executes; it never holds the funds. Honest judgment is made the paying strategy through the structure of the reward, not by an appeal to a higher authority, because there is none. There is no company, no token, and no owner. The protocol charges nothing and rules on the trade, never on the person. It cannot be proven in advance. Its worth is a question for the market, answered by adoption.

1. The problem

Two strangers agree to trade. Each must decide whether the other will perform. If one defects, someone must decide who is owed what. Both decisions, trust before the trade and judgment after it, are made today by a trusted third party.

Every such party is a single point of control. A platform censors, extracts rent, and dies by regulation. A named arbiter can be bribed, coerced, or subpoenaed, and carries the legal liability alone. A rating bureau owns your reputation and rents it back to you. A vote weighted by tokens sells influence to whoever buys the most. And the reputation you build on one platform stays there, and vanishes when the platform does.

The function that has never been decentralized is judgment. Settlement of money became trustless once it was made objective and replicated. Judgment resisted, because a verdict is a matter of preference and there is no sum to check it against. Judgment here stays subjective. The protocol makes honest judgment the paying strategy and removes every party that could be captured.

2. Karma

Karma is a reputation a person earns and carries. It is a single measure of demonstrated trustworthiness that a stranger can read at a glance, without knowing the person's name, location, or past.

It does for trade between strangers what a credential does inside an institution. It certifies that the holder has, over time and at cost, behaved as someone worth dealing with. A degree says a stranger was taught to behave. Karma says a stranger has behaved.

Karma is the product. Dispute resolution is one use of it. Every action on the protocol moves it, up or down: a settled dealing, a vouch that holds, a case judged well add to it; a defection, a broken vouch, a careless verdict take it away. A person accumulates karma through settled conduct and spends its standing wherever trust is priced: buying and selling, borrowing, hiring, joining a community that admits only members above a threshold. A platform that reads karma inherits a user base that is expensive to fake and cheap to filter, and needs no central moderator to do it. The reputation belongs to the person, not the platform, and travels when the person leaves.

Karma is held by no one. It is computed from a public record of signed events, the way a balance is computed from a ledger. Two accounts that praise each other in a closed loop earn nothing, because the measure is trust flowing in from the rest of the graph, and an isolated ring has no inflow.

How far an action moves karma is set by five things, each a curve rather than a fixed amount. The change scales with what the actor put at risk. Each further unit of standing drawn from the same source is worth less than the one before, so no single stake or relationship can be repeated into dominance. The whole is held under a ceiling set by tenure, which no amount of money lifts. A dealing with a fresh and independent party counts for more than another with a familiar one. And each unit of inflowing trust is weighted by the standing and the independence of its source, so trust routed through a captured cluster arrives discounted. Standing earned broadly and slowly compounds. Volume bought quickly does not.

3. The costly signal

A rating is worthless if it costs nothing to give. The reason is older than the protocol.

> Speech is nearly free. A man may say one thing and do another. Action expends scarce time or capital, and reveals, exactly and irreversibly, what the actor valued at the moment of choosing. What he paid for is who he is.

Value is shown in action, never in words. A stated opinion is one more cheap act of speech. A choice that costs something is the only evidence of a preference. So karma moves only on paid-for acts. Completing a trade under bond, standing on a verdict, vouching for a stranger: each costs the actor something he cannot recover, and each is therefore worth recording. A profile description, a self-endorsement, a free rating: each is speech, and moves karma by nothing.

Giving karma costs the giver. To vouch for another account is to stake your own standing on it. If the account you vouched for is later judged a fraud, your stake is cut. A vouch is a demonstrated preference, priced in the one currency that cannot be printed: reputation you spent real time earning.

Endorsement is rationed so it cannot be sprayed. Each further vouch an account gives within a period costs sharply more than the last, so one sincere vouch is cheap and a hundred are not. And because a fraud is charged back along the edges that endorsed it, a careless voucher pays for the account he certified.

4. The price of a signal is time

If the only cost were money, justice would favor the rich, who pay it easily, and the protocol would fall to whoever holds the most. Money is a poor measure of commitment for the same reason it is a

poor measure of a man.

Time is the honest price. No one can buy more of it. A day costs the wealthiest attacker and the poorest honest user exactly the same, and neither can purchase a longer past. Time also cannot be run in parallel inside one identity. A computation that takes a year takes a year on one machine or a million, a fact that can be proven rather than assumed.

Karma is therefore bounded by tenure. Standing accrues only as real time passes with the account in good standing, measured against the timechain, a clock no participant controls. Money enters as one input, held at risk and with sharply diminishing effect, so that capital buys a fast start and never a commanding position. A wealthy attacker reaches the ceiling the buyable inputs allow within a day, then waits years under the same clock as everyone else. The advantage that remains is seniority, not wealth, which is the milder bias and dilutes as the network grows.

5. Sovereignty

The protocol rules on the trade, not on the person. It decides who owes what under an agreement two parties entered freely. It does not decide whether the trade should have been made, whether it was legal, or whether either party is virtuous. Those belong to the individual, who bears the consequences.

This is the third commitment, beside karma and the quorum: sovereignty. There is no supervising authority in the protocol and none above it. Rules hold because participants keep them, and hold only over what participants have agreed to, in the way a network keeps a rule its members chose to run.

Sovereignty extends to the reputation itself. Karma belongs to the account, and an account can be sold, the way a business or a brand is sold, carrying its standing with it. The protocol cannot verify who holds a key, and does not try. That a key's holder is unknowable is what keeps identity sovereign, the same property that lets money be held without permission. What the protocol can do is mark a sharp change in how an account behaves. It reports the change as evolution, not suspicion, because people do change, and it never claims to know whether the hands changed with the conduct. A recently transferred account carries a visible marker of that transfer, and nothing more.

One person may hold many accounts, unlinked, and reveal only what a given dealing requires, because privacy is the power to selectively reveal oneself. And a person may always abandon an account and begin again, at the low standing every new account starts from, carrying none of the old history, good or bad. A fresh start is real, and so is its cost.

6. Escrow

A verdict must move funds or it is only an opinion. The trap is to hand the funds to whoever enforces the verdict, which restores a trusted party at the last step. The protocol splits the problem at the line between what is subjective and what is objective. Judging the dispute is subjective, and is left to the quorum. Checking that the quorum ran correctly is objective, and is left to arithmetic anyone can repeat.

Funds sit in a two-of-two output controlled by the buyer and the seller alone, holding the price and both parties' deposits. Its outcomes are signed in advance. If both agree, at the start or once the trade completes, they close it cooperatively and the quorum is never called. If the trade stalls and neither party opens a dispute, a timeout returns the funds. Once either party opens a dispute, the timeout is void, and the only remaining path is the one the quorum's verdict unlocks.

The quorum publishes a single signature over its verdict. That signature completes exactly one pre-signed outcome, buyer paid, seller paid, or funds returned, and no other.

The quorum attests and never holds. It cannot take the funds, cannot freeze them, and owns no key to the escrow. The two traders are the only signers, so no third party has custody, and the legal exposure that custody carries does not arise.

A verdict can rule only on facts that reach the record. A forged image of a bank transfer, set against a claim that nothing arrived, cannot be judged by anyone. The protocol therefore handles disputes over facts that can be shown: proof of payment, delivery committed to a hash, a signed receipt from a carrier. Claims that rest on a private, reversible rail wait for a proof that can be checked.

7. The quorum

A dispute empanels a quorum: an odd number of accounts, called stewards, drawn from those above a karma threshold, several steps away from both parties in the trust graph, so that no one who knows either trader can sit. The stewards are anonymous to the traders and to each other. They never speak. Each reads the same case file, the evidence and statements both sides submit, stripped of personal detail, and each votes alone.

Within a case the quorum holds all the power, and nothing sits above it. There is no judge to overrule it and no court to appeal to. The absence is deliberate. Every arbitration system before this one ends at an authority that can be captured, and the search for a final, incorruptible authority is the search that has always failed. The protocol puts in that place a crowd that no one can find, bribe by name, or coordinate, and takes the crowd's verdict as final. Finality is itself the value. A dispute settled once, in days, is worth more than one litigated for years through a system backed by force. An appeal path was considered and left out, because any right to reopen a case invites the patient and the rich to grind the poor into surrender.

Each side is given time to state its case and submit evidence, and the quorum decides on that alone. Some people argue better than others. The protocol does not correct for this and does not pretend the world is otherwise. It weighs the case as presented, as any reader of evidence must.

8. Sortition

The quorum must be unpredictable before it is drawn and unknown after, or it can be bribed, threatened, or packed. Public randomness from an outside committee would restore a party to trust. The randomness is taken instead from the protocol's own activity.

Every account carries a seed that changes as the account acts. Each action folds a committed trace of itself into the seed at the moment it happens, on an append-only log, so the trace cannot be altered or chosen after the fact. The seed is the account's own irreversible history, reduced to entropy.

When a quorum is due, the seeds of a broad set of active accounts are combined into one value. That value alone is still not safe, because the last account to act before the cutoff could try many actions and keep the one that steers the draw. So the combined value is passed through a delay function: a computation that takes a fixed span of wall-clock time to run and an instant to check, and that no amount of parallel hardware can hurry. No one, not even the last actor, can learn the result in time to bias it. Its output is the seed for the draw.

Each eligible account then checks, privately, whether it was selected, by evaluating a verifiable function of the seed under its own key. Selection proves itself and reveals nothing about the others.

An account learns only that it was drawn, and can prove that fact without naming itself. The set that decides a case is a group of pseudonyms that exists for that case alone, unknown to the traders, to each other, and to the operators of any relay. There is no list to subpoena and no colleague to conspire with.

No randomness of this kind is perfectly unbiasable, and this is not. The delay function rests on a bound on how fast the fastest hardware can run it, and the draw can be skewed only by an actor who both dominates the activity in the window and holds a decisive speed advantage over everyone else. The design drives that margin small enough that skewing a draw costs more than winning the dispute it would decide. That is the standard the protocol holds, in place of a proof of the impossible.

Two further rules keep the draw clean. Karma is frozen the moment a draw opens, so no account can lift its own standing to improve its odds once selection is under way. And the share any single trust cluster can hold, both of the eligible set and of the panel actually drawn, is capped, so a group that certifies itself cannot pack a quorum however large it grows.

9. The verdict

The quorum does not find the truth. No crowd can, and the protocol makes no such claim. It aligns the reward with honest reporting, and lets the honest report win.

Only the stewards who land with the majority are paid. Those in the minority are not. A steward who votes without reading gambles against everyone who read, so the paying move is to read the evidence, judge it, and expect others who read to judge as he does. Payment for agreement, with a penalty for lazy or dishonest votes, makes honest attention the profitable strategy.

Plain majority reward has a known failure. On a small stake, a steward may reason that the safe vote is whatever the crowd will pick, and vote to predict the crowd instead of to weigh the case. The protocol closes this by asking each steward two things: the verdict, and a guess at how the others will vote. The answer that beats its own predicted popularity wins. A steward with private reason to judge against the obvious answer is rewarded for it when the reason is sound, which is the case a plain majority gets wrong. Votes are sealed until all are cast, and there is no deliberation, so there is no channel to coordinate through and no shared error to amplify.

Three defenses hold the vote honest under pressure. Ballots are sealed so that only the totals ever open, never any single vote, and a steward who is leaned on can quietly change a vote already cast, so a briber can never confirm how anyone voted and a payment offered for a vote has nothing left to condition on. The panel grows with the stake, a few stewards for a small claim and many for a large one, so the pay for reading always clears the cost of reading and no large sum ever rests on a handful. And a steward serves a batch of cases at once and is scored across all of them, where honest reading beats any lazy or collusive rule whatever the others do, a guarantee that holds only across many cases and never on a single one.

The penalty for a wrong verdict rises with how many stewards err together on the same case. An isolated mistake costs little. A bloc that moves the same wrong way pays toward the whole of its stake, because strangers sworn to silence and drawn apart should not fail in unison, and when they do, coordination is the likeliest cause. Punishing correlated error hardest turns the profit of collusion negative before a bloc can form.

The ballot is guilty, not guilty, or unresolvable. The last returns the funds. It exists because the most common hard case is ambiguity, not deceit, and forcing a call would reward whoever engineered the ambiguity. On a grave case, and only when a majority of the quorum calls for it and an added bond is posted, the stewards may put written questions that both parties must answer, so

a packet built to mislead can be probed. The quorum ranks its own questions with no moderator: each steward spends a fixed budget of weight across the questions at a rising cost per question, so only what matters draws weight, and a question that both sides of a divided quorum press rises above the merely popular one. The few at the top go to the parties. The ordinary dispute carries no such round and stays cheap to settle.

A steward is judged in turn. A verdict is an action, recorded, and later verdicts can bear on it. A steward whose verdicts are repeatedly overturned by the quorums that follow loses karma and is drawn less often. A steward who serves well gains standing. No authority reviews the stewards. Their record reviews them, as every account's record governs its standing. This is how the protocol carries personal responsibility into the one role that decides other people's disputes, without appointing anyone to watch the watchers.

10. Economics

Both traders post their deposits and the dispute fee into the escrow when the trade is set, so opening a dispute costs no new money and a poor honest party is never priced out of defending itself. Deposits scale down with karma, so an established account posts little and a fresh account posts much, and scale up with the amount in dispute, so corrupting a case always costs more than the case is worth.

On resolution, the losing side's deposit and the bonds of stewards who voted against the outcome pay the stewards who voted with it. Attacking and losing burns the attacker's capital. The winner is made whole.

The protocol takes nothing. Every fee flows to the people who did the work of judging. There is no issuer to pay, because there is none. That there is no earning model is a requirement. Anyone who could profit from the court could sell it, and a court that can be sold is a court that will be. A fee-taking company is a standing target for a subpoena and a bribe. A token is judicial power for sale, and turns the court into the property of whoever buys the most. The protocol has neither, and is funded the way shared infrastructure is funded, by grants and by the people who want it to exist. Neutrality is the product. An owner would be the flaw.

11. Sybil resistance

The security of the protocol is the cost of faking karma. That cost must be real, and cheap neither to a patient attacker nor to a rich one.

Creating an account costs an irreversible sacrifice, uniform for everyone, so that a thousand accounts cost a thousand times one and no volume discount exists. Karma is minted only by settled dealings with distinct counterparties, never by self-endorsement, so a ring trading with itself earns nothing. Standing accrues slowly, each further unit from the same source worth less than the last, and decays unless renewed by fresh conduct, so a burst of early activity bleeds away and no account can bank a fortune of karma before the network is populated. Weight flows in from the established graph, so an account is only as trusted as the independent accounts that vouch for it.

One problem stays open and is stated plainly. Telling distinct persons apart without an identity document is unsolved at scale, and the protocol's resistance to an attacker who can cheaply create real-seeming humans is only as strong as the best answer to that problem. The protocol raises the cost of an attack. It does not claim to make one impossible.

12. Adoption

The protocol cannot be tested in a laboratory. Trust between strangers is a network good, worth little to the first user and much to the millionth, and no simulation produces the thing being tested, which is a large population of real people with real reputations at stake. It will be adopted gradually if it is useful, or not at all. This is the condition faced by every protocol that replaced a trusted party with a network, and no engineering removes it.

The path is to earn karma inside systems that already exist. A marketplace, an exchange, or a community that adopts the protocol gives its users a reputation they can carry out, and gains a filter against fraud that no central moderator can match. Where a costly reputation graph already exists, the protocol can read it, so that standing earned elsewhere seeds standing here, and the first quorum is not conjured from nothing. Chronicle Network, an open and ownerless graph of costly, signed reputation built by Luka Dover, is one such source and a companion to this work: it holds the standing, this protocol holds the judgment. One key is a participant in both, so a verdict here can be a marked event there, and a steward who judges well is credited on a graph wider than this one. Where Chronicle already carries a person's standing, the cold start is borrowed rather than bootstrapped. Chronicle Network, an open graph of costly, signed reputation, is one such source and one natural partner: it supplies the standing, this protocol supplies the judgment.

13. Beyond disputes

The quorum is a general instrument. It decides any question that needs an impartial answer and has evidence to decide on, not only who is owed money after a trade. A group can put a decision to it. A community can settle a rule with it. Any process that today relies on a trusted counter, a trusted moderator, or a trusted authority can be handed to a crowd that cannot be found or bought.

A question outside trade has no losing deposit to pay the stewards, so it funds itself. Whoever poses it posts a bounty that becomes the reward, and each steward posts a bond of standing to take part. Many such questions have no outcome to check a vote against, so the stewards are scored on one another: each submits a sealed answer and a guess at how the others will answer, and the answer that beats its own predicted support is paid, while the rest is slashed and given to those who were right. The panel is settled against itself, with no outside loser and no owner to subsidize it. Every such decision is recorded as an action, so the standing a person earns in trade is the standing they stake in judgment, and one reputation follows them across everything the protocol touches.

The furthest form of this is a way to reach a binding collective decision that no one can quietly corrupt, at a cost of corruption far above the value of any single outcome. A vote fails where the count, the roll, or the authority can be captured. A verdict drawn from an unfindable, bonded, uncoordinated crowd removes each of those targets. Whether it replaces anything is for the people who use it to decide. The protocol only makes it available.

14. Limits

The protocol rests on reasoning about how people act, not on evidence that it works, and it keeps the two apart. It cannot be proven in advance, and a correlation in early data would not prove it either. What can be shown is where it must hold and where it cannot.

Four limits survive every part of the design and are named rather than buried. An account's signing key can be bribed, and no sealed ballot stops its owner from selling the vote outright. Anonymous stewards who find a channel outside the protocol can coordinate, and the reward for

honest reporting weakens when they do; what stands against this is that the stewards know nothing of each other, so coordinating means building a separate conspiracy among people the system keeps apart, against the very system their standing is worth something in. Telling distinct persons apart without an identity document is unsolved, and the defense against mass-produced identities is only as strong as the answer to that problem. And the final defense against an attacker who has bought the whole graph is that the graph is expensive to buy, not that it cannot be.

The protocol therefore does not claim to be unstoppable. It claims to be the first way to resolve disputes and carry reputation with no single party to bribe, subpoena, or switch off, and to make the cost of capturing it exceed the value it protects, for stakes kept within that bound while the network is young.

15. Conclusion

Trade between strangers has always rested on someone trusted to vouch and someone trusted to judge, and every such someone has been a point of failure. This protocol earns the vouch from costly action and draws the judgment from a crowd that no one can find, holds the funds with the traders alone, and keeps no power for itself. The pieces are known: reputation from real conduct, selection at random, sealed votes, an oracle over an escrow, a clock no one owns. What has not been built is their assembly into a whole that no one owns and no one can capture. The result belongs to the public domain, to use, to build on, and to improve, without permission and without an owner, because a thing meant for everyone must be owned by no one.

References

To be finalized. Prior art and grounding by theme: reputation and sybil resistance; sortition and verifiable randomness (delay functions, verifiable random functions); the mechanism design of honest reporting (peer prediction, the surprisingly-popular answer); escrow without custody; the costly-signal foundation in the theory of human action; portable reputation on an open graph (Chronicle Network, Luka Dover / Cartographers Guild); privacy as selective disclosure (Eric Hughes, 1993). A full annotated list is maintained separately.
