Proof of WiseWork:
A Peer-to-Peer System for Truth Consensus

Fatih Dinc
fatdinhero@gmail.com
Pforzheim, Germany

Version 2.0 – April 2026
DOI: 10.5281/zenodo.19642292
OTS: poisv.com/verification

Abstract

A purely distributed form of truth discovery would allow claims to be validated directly between independent observers without going through a trusted authority. Digital signatures provide part of the solution, but the main benefits are lost if a trusted third party is still required to assess the correctness of transmitted information. We propose a solution to the distributed-truth problem using a peer-to-peer network of wisdom validators. The network computes a composite WiseScore from truth, context, relevance, and ethical compliance for each proposed block, and additionally tests a non-collusion statistic grounded in the Meta-Bell Theory [7]. Accepted information units are permanently anchored on a DAG-ordered chain, forming a globally available, tamper-evident knowledge base. The system is secure as long as honest validators collectively perform more independent semantic work than any coordinated group of attackers.

1 Introduction

Consensus on the Internet today relies almost exclusively on trusted intermediaries. Financial platforms agree on account balances through central ledgers, social platforms agree on permissible content through central moderators, and artificial intelligence systems produce outputs whose correctness users cannot independently verify. This model works for much of everyday traffic but suffers from the well-known weaknesses of the trust paradigm. The costs of trust rise with the importance of the decision, and a baseline rate of errors and manipulation is accepted as inevitable.

Existing decentralised consensus mechanisms solve a narrower part of the problem. Proof of Work agrees on an ordering of transactions by expending computational effort. Proof of Stake achieves the same ordering by locking capital as collateral. Neither mechanism makes any verifiable claim about whether the content of the ordered information is true, contextually appropriate, relevant, or ethically sound. For pure payments this suffices; for the information age it does not.

What is needed is a consensus mechanism in which not raw computational effort or capital, but demonstrably wise work is the scarce resource. We call such a mechanism Proof of WiseWork (PoWW). This paper makes four contributions. We define a composite score over information units that integrates truth, context, relevance, and ethical compliance. We show how a statistic derived from the Meta-Bell Theory can be used to distinguish colluding validators from independent ones. We describe how accepted information units are permanently anchored on-chain, forming a publicly

accessible, tamper-evident knowledge base. We derive the probability of an attacker succeeding and show that it falls exponentially in two parameters.

2 Information Units

We define an information unit as a four-tuple $i = (v, c, r, e)$ with a truth candidate $v \in [0, 1]$, a context weight c in the non-negative reals, a relevance factor r in the non-negative reals, and an ethical compliance value $e \in [0, 1]$. The four components may be produced by different procedures, provided those procedures are deterministic and reproducible by all nodes.

The problem is that a single information unit allows no reliable inference about the collective truth of a set of claims. A high truth score may be invalidated by context shift; high relevance may rest on false assumptions; a favourable context may highlight irrelevant content; and a technically true claim may violate ethical principles. Only the combination of all four dimensions yields a robust indicator. Analogously to the signature chain of an electronic coin, an information unit does not carry its legitimacy in itself alone but in its relation to other units.

The truth value v can be defined on four levels. At the empirical level, $v_i =$ validated evidence/total evidence. At the physical level, $v_i = 1 - |x_{\text{measured}} - x_{\text{theory}}|/x_{\text{theory}}$. At the quantum-probabilistic level, $v_i = |\psi_i|^2$, where ψ_i is the amplitude for event i. At the information-theoretic level, $v_i = 1 - H(i)/H_{\max}$, where $H(i)$ is the Shannon entropy. Analogously, the ethical compliance value e can be defined on four levels: as degree of conformity with codified legal norms, as degree of fulfilment of deontological rules, as a consequentialist net-utility assessment, or as conformity with a contextual value base. The choice of level depends on the application, and all definitions are compatible with the common acceptance rule.

3 DAG Ordering Layer

Validators work in parallel on semantic tasks that cannot be reduced to a linear order. A classical blockchain consensus would discard a large proportion of this parallel work as orphaned blocks. We therefore choose a directed acyclic graph structure as the ordering layer, specifically the GHOSTDAG protocol by Sompolinsky, Wyborski, and Zohar [5]. In this protocol, each new block references all observable tips as parent blocks, and a k-cluster algorithm subsequently induces a total order over the included blocks. The ordering layer is content-agnostic; semantic validation is performed in a layer placed on top of it, described in Section 5.

4 Proof-of-WiseWork

Let I be the set of all information units of a block. We assign each unit four normalised quantities. The normalised truth is

$$T(i) = \frac{\exp(\alpha \cdot v_i)}{\sum_{j \in I} \exp(\alpha \cdot v_j)} \tag{1}$$

with sharpness parameter $\alpha > 0$. The choice of softmax normalisation follows from the maximum-entropy principle under a linear constraint on the expected truth value; it is the unique map into the probability simplex that is positivity-preserving, invariant under additive shifts, and consistent with minimal Kullback–Leibler projection. The context weight is

$$C(i) = \frac{c_i}{\sum_{j \in I} c_j}. \tag{2}$$

The relevance is

$$R(i) = \log(1 + r_i). \tag{3}$$

The logarithmic dampening ensures that relevance inflation by an attacker enters the total score only sub-linearly. The fourth component is the ethical compliance value,

$$E(i) = e_i, \tag{4}$$

as a normalised value measuring the information unit's conformity with the explicitly chosen value base of the application domain. The composite WiseScore of a unit is the product of all four components,

$$W(i) = T(i) \cdot C(i) \cdot R(i) \cdot E(i). \tag{5}$$

The multiplicative form follows from the requirement that a unit must score well on all four dimensions to make a significant contribution. Additive forms would allow one dimension to compensate for others, destroying the core idea. An ethically problematic but technically true result thus receives a score near zero regardless of the other components. The aggregate block score is the arithmetic mean,

$$\text{PoWW} = \frac{1}{|I|} \sum_{i \in I} W(i). \tag{6}$$

The acceptance rule additionally requires a non-collusion condition grounded in the Meta-Bell Theory [7]. We treat the validators as measuring apparatuses and their outputs as measurement results. From the empirical correlation structure of the outputs we derive a statistic Ψ that measures the deviation of the observed joint distribution from the set of all local hidden-variable models. In the special case of two validators with two measurement settings each and binary outcomes, Ψ reduces to the normalised deviation of the classical CHSH quantity from the bound of two (Lemma 1, proof in the appendix). We interpret $\Psi > 0$ as statistical proof that the observed agreement of the validators cannot be explained by a shared hidden variable. A block is accepted if and only if

$$\text{PoWW} \geq \theta \quad \text{and} \quad \Psi \geq \Psi_{\min}. \tag{7}$$

The role of Ψ in the system can be stated precisely: while the WiseScore is the operative aggregate metric that enters consensus, Ψ is the mathematical proof of the quality of the underlying validator work. The product of both quantities forms the complete acceptance signature of a block.

5 Network

The steps of the network are as follows.

1. New claim units are sent to all validators.
2. Each validator processes the unit in a Reception Layer against local admission rules, recording timestamp and signature.
3. Each validator computes T, C, R, E, and W for the unit in a Comprehension Layer and records its contribution in the current block proposal.
4. A validator wishing to propose a block generates a zero-knowledge proof of the correctness of the computation in a Cognition-Proof Layer, using the Nexus zkVM [6].
5. The block together with the proof is broadcast to all validators.

6. Validators accept the block if all units are valid, the zk-proof confirms the computation, and acceptance rule (7) is satisfied. Accepted blocks are added to the DAG as witnesses for the next block round.

Validators always consider the DAG with the highest cumulative WiseScore as the correct one. If two validators receive different next blocks, they work on both branches and prefer the one whose score reaches the next acceptance level first. New claims do not need to reach all validators immediately; as long as they reach many, they enter a block before long.

6 Incentive Structure

The first transaction entry of a block is a special credit assigning new native currency units to the validator who proposed the block. This rewards validators for contributing genuine information units to the network and distributes the currency incrementally without a central issuer. Participants can additionally pay fees for the validation of specific claims; these fees also flow to the proposing validator.

A rational attacker with sufficient resources to circumvent the acceptance rule faces a simple cost–benefit calculation. He can use his resources to behave like an honest validator and collect the regular block reward, or he can expend them to manipulate the history. The second variant requires not only that his block exceeds the score threshold, but also that his validator outputs reach the Ψ threshold. The second condition cannot be met by coordination, because any coordination depresses the statistic below the threshold. The attacker would therefore be forced to bring a cohort of genuinely independent validators under his control, which is incompatible with his original intention of coordination. He therefore finds it more profitable to follow the rules.

7 On-Chain Knowledge Base and Storage Recovery

Accepted information units are permanently anchored on-chain together with their full score and zk-proof, forming a globally available, tamper-evident knowledge base in which every legitimised claim is stored with a cryptographically attested quality certificate. Every participant can query this base to retrieve for any given claim its current score, contextual embedding, evidential basis, and ethical assessment. This fundamentally distinguishes PoWW from classical blockchains, which digitalise only ownership relationships, whereas PoWW digitalises validated knowledge itself and makes it publicly accessible.

Once the WiseScore of an information unit has been confirmed sufficiently often, the underlying raw validator outputs can be discarded without breaking the block hash. Validator outputs are hashed in a Merkle tree; only the root enters the block header. Old blocks can thus be reduced to the header plus the hash path to the currently referenced units. A block header of approximately 200 bytes, a zk-proof of approximately 200 kilobytes, and a small index per day suffice to require approximately 6 gigabytes per year at a block rate of one block per second. With current storage technology this is unproblematic, since the raw data can be reconstructed from the sources if needed.

8 Simplified Verification

It is possible to verify the acceptance rule without running a full validator node. A client needs only to hold a copy of the block headers of the currently longest chain with the highest cumulative score, which it can obtain by querying selected validators. It cannot verify the semantic computation

itself, but relies on the zk-proof attached to the block and on the fact that subsequent blocks confirm acceptance.

Verification is reliable as long as honest validators control the network. Businesses that frequently receive information units will nonetheless operate their own validators to obtain independent security and faster verification. For end users, the light client suffices as long as they can rely on the validator community as a whole. Light clients may additionally apply subjective filters, for instance only displaying claims validated by a specific ethics committee, without imposing these filters on the network as a whole.

9 Combining and Splitting Claims

While claims could be handled individually, it would be unwieldy to require a separate validation for each sub-claim of a complex assertion. A claim therefore contains multiple input and output edges that split or bundle atomic sub-claims. Normally there is one input from a larger predecessor claim or multiple inputs combining smaller sub-claims, and at most two outputs, one for the accepted sub-claim and one for the remaining uncertainty residual. Fan-out, where a claim references multiple predecessors and those reference further claims, is no problem here, since the complete history of a claim never needs to be reconstructed.

10 Privacy

The traditional model achieves a degree of privacy by restricting access to the parties involved and a trusted intermediary. The public announcement of all information units precludes this method, but privacy can be maintained at a different point. The zkVM witness contains the raw evidence sources or model weights and remains private; only the result of the acceptance rule and the numerical score are made public. The publicly visible data thus correspond to the times and sizes of individual operations, as a stock exchange makes these public without revealing the identities of the parties involved.

As an additional shield, a new key pair should be used for each claim submission so that submissions cannot be linked to a common owner. Some linking is unavoidable with multi-input claims, since these necessarily reveal that their inputs belong to the same owner. The Ψ test operates exclusively on the correlation structure of the public validator outputs and requires no content data; it thus provides a non-collusion signature without itself disclosing additional information.

11 Calculations

We consider the scenario of an attacker who attempts to generate an alternative PoWW chain with a higher cumulative score than the honest network. Even if he achieves this, it does not open the system to arbitrary changes, such as creating value from nothing or appropriating foreign information. Validators do not accept an invalid block as payment, and honest validators will never include a block that violates the acceptance rule in the chain. An attacker can only attempt to retract one of his own recent information units.

The race between the honest chain and an attacker chain can be characterised as a random walk. The success event is the extension of the honest chain by one block, growing the lead by $+1$, and the failure event is the extension of the attacker chain by one block, reducing the gap by -1. The probability that an attacker catches up from a given deficit is analogous to a Gambler's Ruin problem [8]. We set

p = probability that an honest validator finds the next block
q = probability that the attacker finds the next block
q_z = probability that the attacker catches up from a deficit of z blocks

$$q_z = 1 \text{ if } p \leq q, \qquad q_z = (q/p)^z \text{ if } p > q.$$

Under the assumption $p > q$, the probability falls exponentially as the number of blocks the attacker must catch up grows.

We now consider how long the recipient of a new information unit must wait before being sufficiently certain that the sender can no longer alter it. The attacker, once the unit is sent, begins working in secret on a parallel chain containing an alternative version of his claim. The recipient waits until the unit has been included in a block and z blocks have been appended to it. He does not know the exact progress of the attacker, but under the assumption that the honest blocks required the average expected time per block, the attacker's potential progress is a Poisson distribution with expected value $\lambda = z \cdot (q/p)$.

To obtain the probability that the attacker could still catch up now, we multiply the Poisson density for each amount of progress he could have made by the probability that he could catch up from that point, and rearrange to avoid summing the infinite tail of the distribution:

$$1 - \sum_{k=0}^{z} \frac{\lambda^k e^{-\lambda}}{k!} \cdot \left(1 - \left(\frac{q}{p}\right)^{z-k}\right).$$

Converting to C code:

```
#include <math.h>
double AttackerSuccessProbability(double q, int z)
{
    double p = 1.0 - q;
    double lambda = z * (q / p);
    double sum = 1.0;
    int i, k;
    for (k = 0; k <= z; k++)
    {
        double poisson = exp(-lambda);
        for (i = 1; i <= k; i++)
            poisson *= lambda / i;
        sum -= poisson * (1 - pow(q / p, z - k));
    }
    return sum;
}
```

Running some results, we can see the probability drop off exponentially with z:

$q = 0.1$	
$z = 0$	$P = 1.0000000$
$z = 1$	$P = 0.2045873$
$z = 2$	$P = 0.0509779$
$z = 3$	$P = 0.0131722$
$z = 4$	$P = 0.0034552$
$z = 5$	$P = 0.0009137$
$z = 6$	$P = 0.0002428$
$z = 7$	$P = 0.0000647$
$z = 8$	$P = 0.0000173$
$z = 9$	$P = 0.0000046$
$z = 10$	$P = 0.0000012$

$q = 0.3$	
$z = 0$	$P = 1.0000000$
$z = 5$	$P = 0.1773523$
$z = 10$	$P = 0.0416605$
$z = 15$	$P = 0.0101008$
$z = 20$	$P = 0.0024804$
$z = 25$	$P = 0.0006132$
$z = 30$	$P = 0.0001522$
$z = 35$	$P = 0.0000379$
$z = 40$	$P = 0.0000095$
$z = 45$	$P = 0.0000024$
$z = 50$	$P = 0.0000006$

Solving for $P < 0.1\%$:

q	z
0.10	5
0.15	8
0.20	11
0.25	15
0.30	24
0.35	41
0.40	89
0.45	340

The preceding derivation is the unweighted score bound. It alone suffices to furnish PoWW with the security of a classical Nakamoto consensus. The Meta-Bell statistic adds a second, independent bound. When k validators are coordinated by a shared hidden variable, their joint output factorises over that variable, and the expected value of Ψ under collusion is zero. Hoeffding's inequality yields:

$$P_{\Psi}(k) \leq \exp\left(-2k \cdot \frac{\Psi_{\min}^2}{\Delta_{\text{crit}}^2}\right).$$

The joint attacker success probability is the product of both bounds. The score bound weakens as q increases, but the Ψ bound becomes exponentially stronger precisely because k grows with q. An attacker with $q = 0.49$ and six blocks of deficit with one hundred validators per block achieves, according to our computations, a joint success probability below 10^{-23}. The attacker success probability thus falls exponentially in the size of the required collusive cohort.

12 Conclusion

We have proposed a system for distributed truth validation that does not rely on a trusted authority. We began with the standard framework of digital signatures, which provides strong control over the origin of a claim but is incomplete without a mechanism for assessing content. To solve this, we proposed a peer-to-peer network that computes a composite WiseScore from truth, context, relevance, and ethical compliance, and tests a Meta-Bell non-collusion statistic to statistically rule out collusion. Accepted units are permanently anchored on-chain, forming a publicly accessible,

tamper-evident knowledge base. The network is robust in its unstructured simplicity. Validators work largely independently and require no identification, since messages are not routed to any specific place and need only be delivered on a best-effort basis. They vote with their semantic work, expressing their acceptance of valid blocks by continuing to build on them and their rejection of invalid blocks by withholding that work. Any needed rule and incentive can be enforced through this consensus mechanism.

References

[1] S. Nakamoto, *Bitcoin: A Peer-to-Peer Electronic Cash System*, 2008.

[2] J. S. Bell, *On the Einstein Podolsky Rosen Paradox*, Physics 1(3):195–200, 1964.

[3] J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, *Proposed Experiment to Test Local Hidden-Variable Theories*, Physical Review Letters 23:880–884, 1969.

[4] C. E. Shannon, *A Mathematical Theory of Communication*, Bell System Technical Journal 27:379–423, 1948.

[5] Y. Sompolinsky, S. Wyborski, A. Zohar, *PHANTOM GHOSTDAG: A Scalable Generalisation of Nakamoto Consensus*, AFT 2021.

[6] Nexus Labs, *Nexus zkVM: Enabling Verifiable Computation*, 2024.

[7] F. Dinc, *Meta-Bell Theory: A Measure-Theoretic Extension of Bell's Inequalities*, v1.0, SHA-256: 062e290009f6b7339e9a8b522ce1d94d9021d109d8e4bc41210d1f3dda053a3b, 2026.

[8] W. Feller, *An Introduction to Probability Theory and Its Applications*, 1957.