KLEROS

kleros.io

25 September 2026

# The Role of Identity

## in Decentralized Dispute Resolution Protocols

**William George**
Research Lead @ Kleros

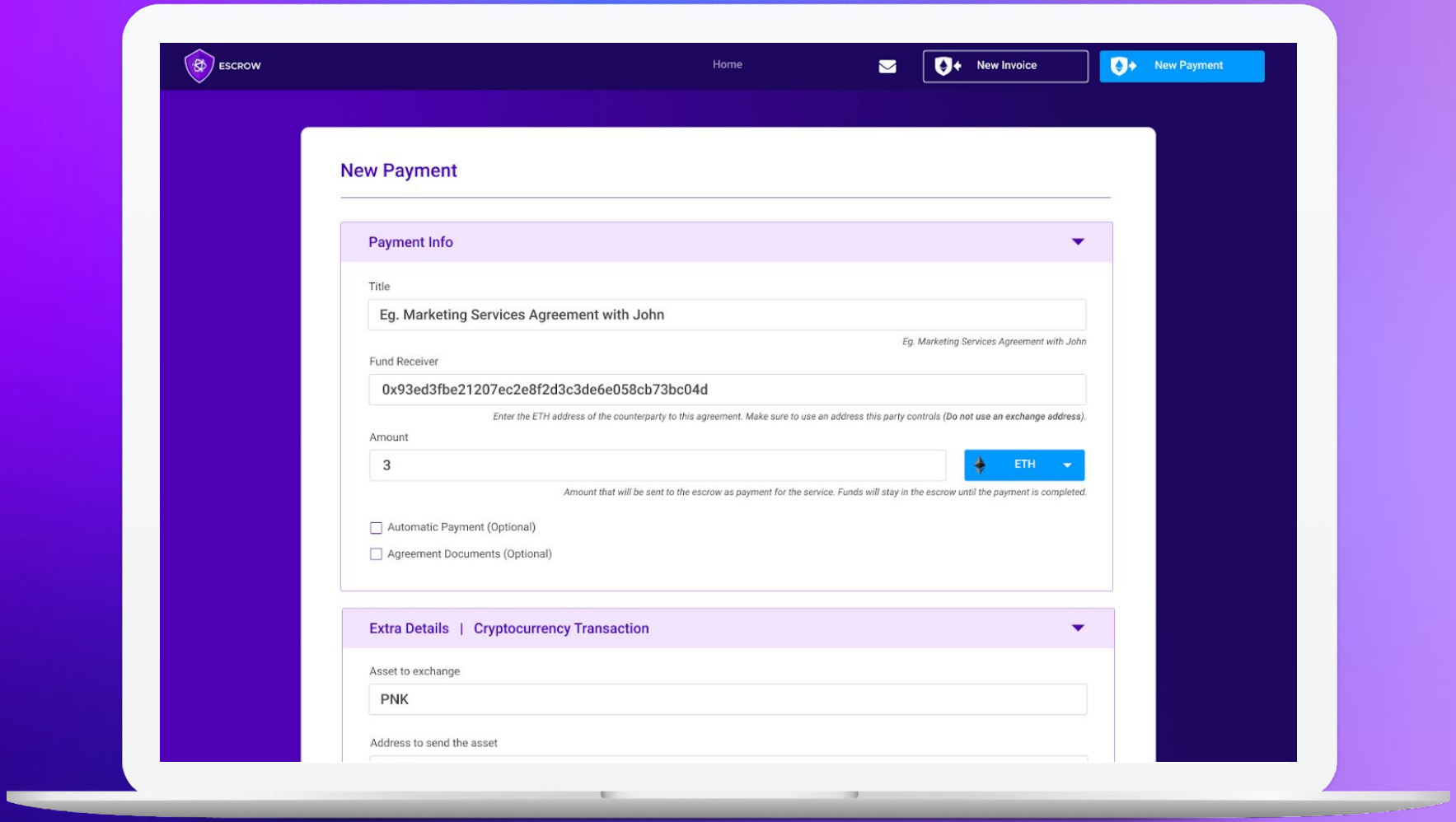
@kleros_io

HOSTED BY THOMSON REUTERS INCUBATOR

European Commission | Winner of the Blockchains for Social Good Prize from the European Union's Horizon 2020 research and innovation programme

bpifrance

BLOCKCHAINGOV


ESCROW
Home
New Invoice
New Payment
New Payment
Payment Info
Title
Eg. Marketing Services Agreement with John
Eg. Marketing Services Agreement with John
Fund Receiver
0x93ed3fbe21207ec2e8f2d3c3de6e058cb73bc04d
Enter the ETH address of the counterparty to this agreement. Make sure to use an address this party controls (Do not use an exchange address).
Amount
3
ETH
Amount that will be sent to the escrow as payment for the service. Funds will stay in the escrow until the payment is completed.
Automatic Payment (Optional)
Agreement Documents (Optional)
Extra Details | Cryptocurrency Transaction
Asset to exchange
PNK
Address to send the asset


KLEROS


KLEROS

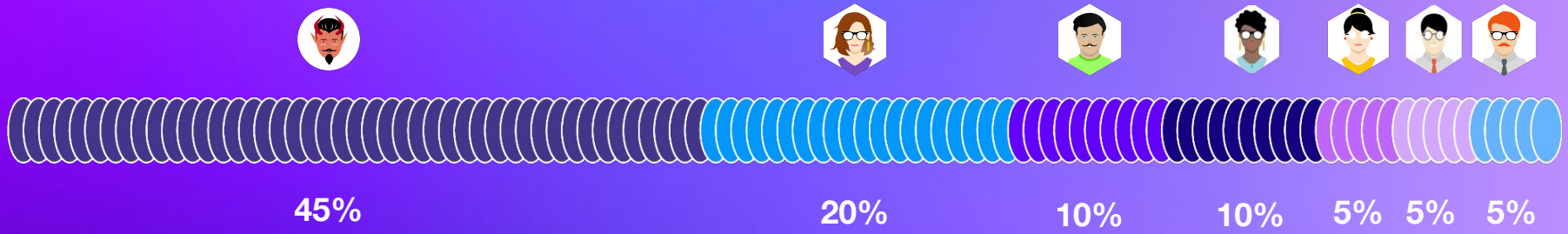

VS





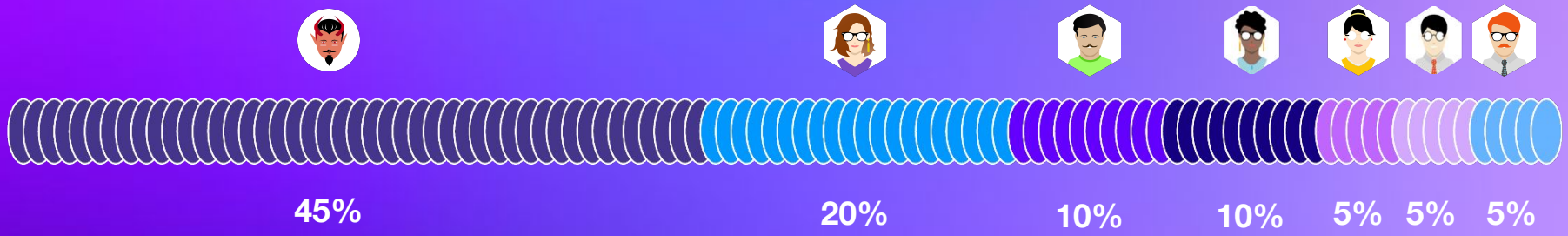


# Token Weighted Drawing




**Panel:** 

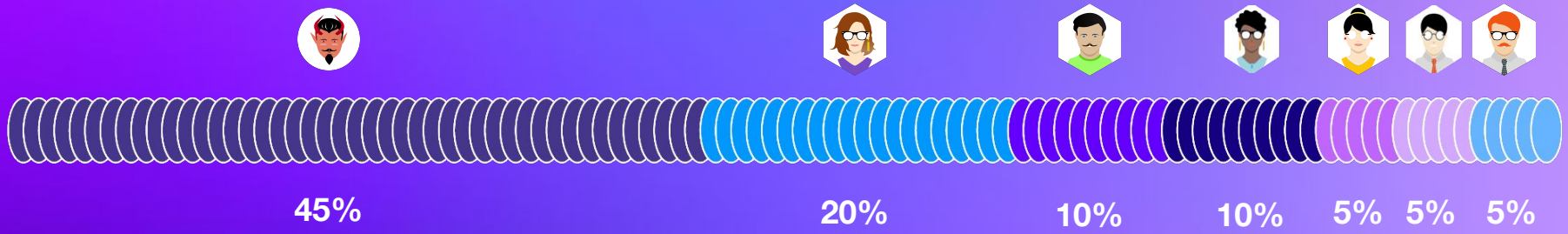
# Token Weighted Drawing




**Panel:**



# Token Weighted Drawing




**Panel:**



45% chance of another



55% chance of someone else

42.5% overall chance of getting >= 2  ‘s

# Kleros 1.0 – **Security Model**

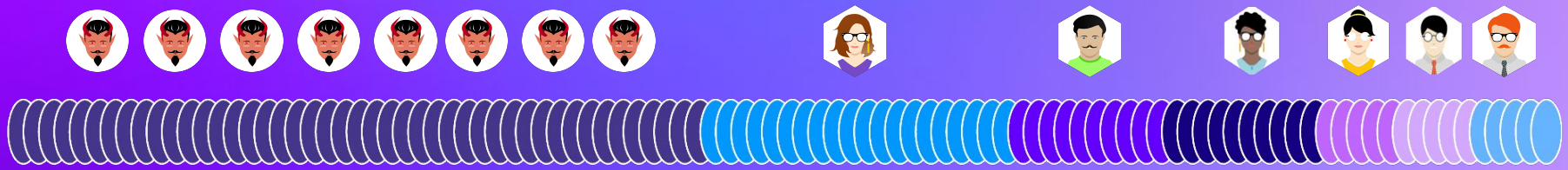
> No participant (or colluding group) has 50% of the stake

# Lots of Attackers: Limited identity tools




Paul Steiner, New Yorker, 1993

# Lots of Attackers: Limited identity tools



45% total divided between many small accounts

20%

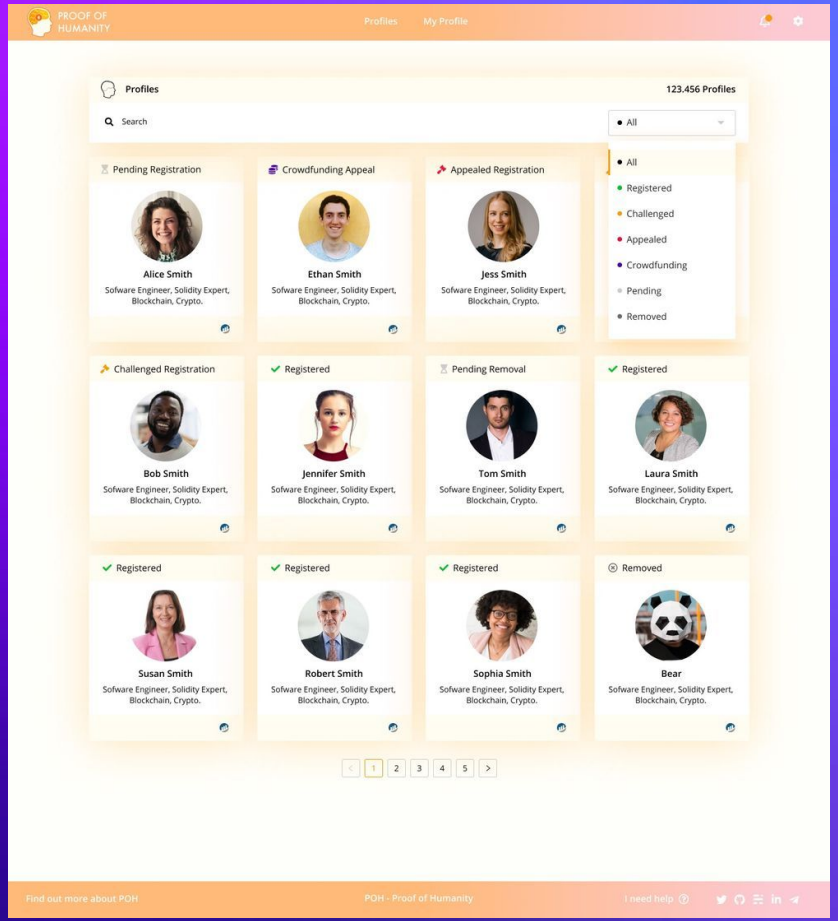
10%

10%

5%

5%

5%



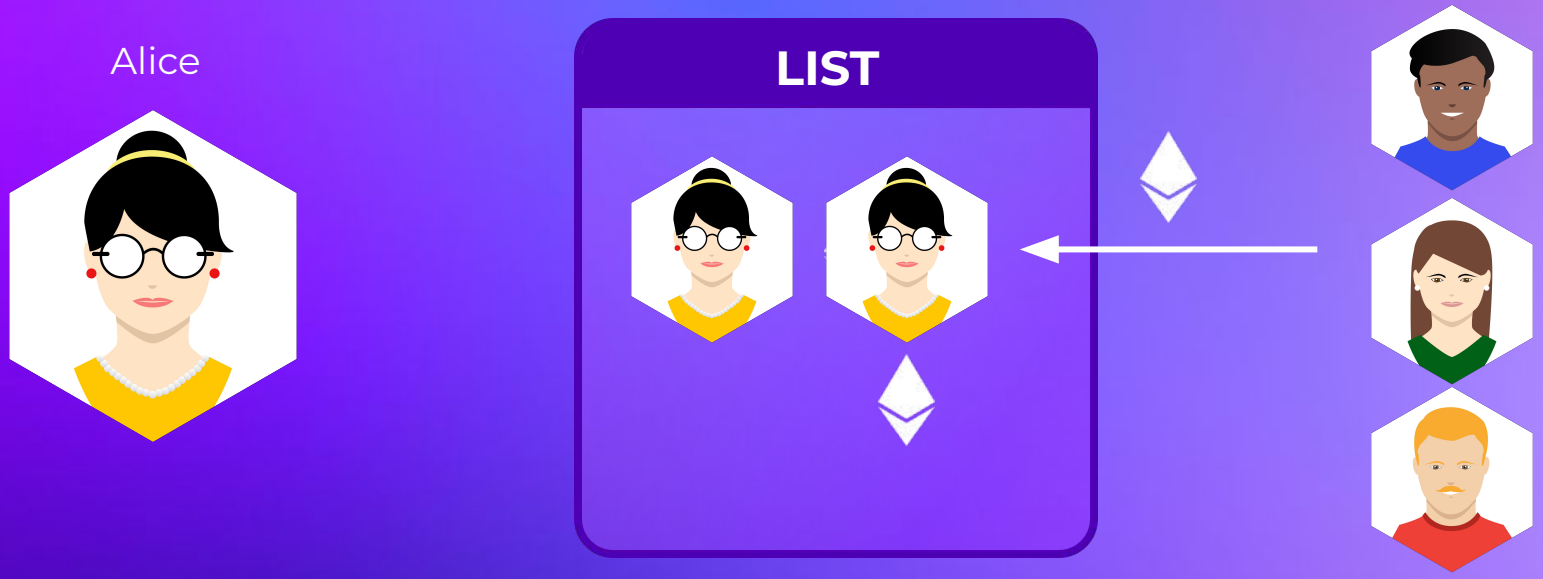



# PROOF OF HUMANITY

- ➔ Proof of personhood/Sybil resistance tool
  - ◆ Alternative to Worldcoin
- ➔ Curated list, if dispute over eligibility -> Kleros dispute




# User Submits Profile




# Challenge



Jury



Submitter



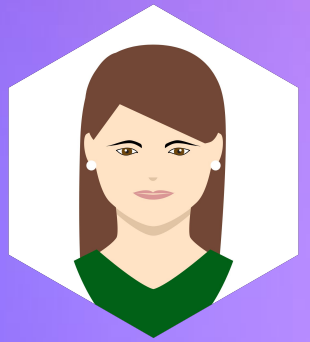


Challenger



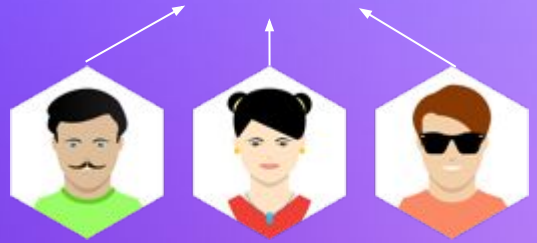




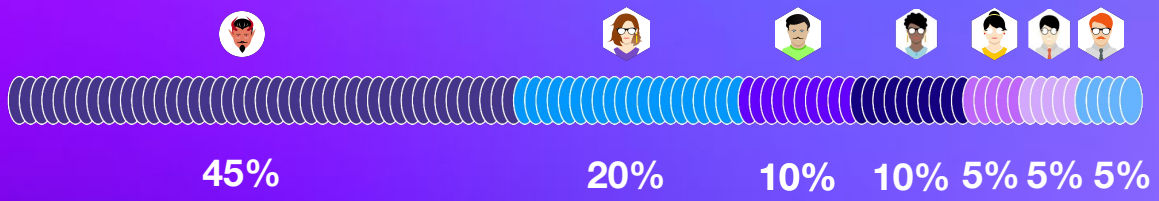



REJECTED

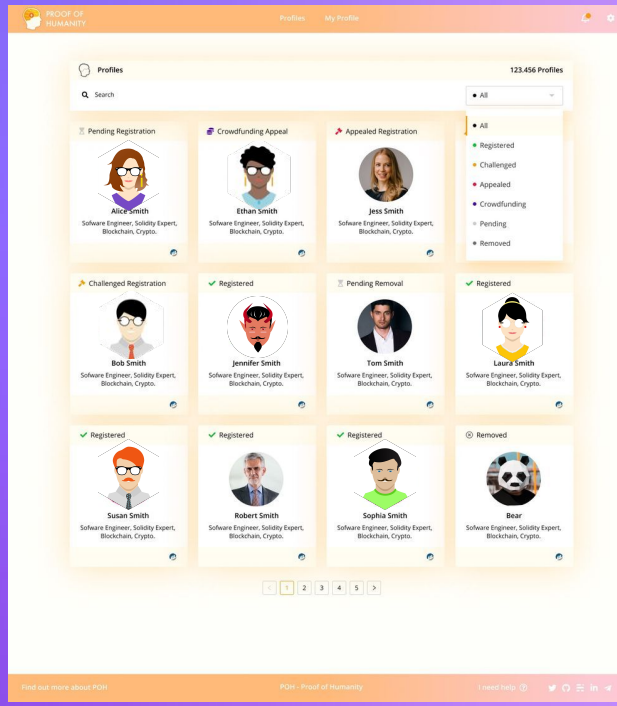




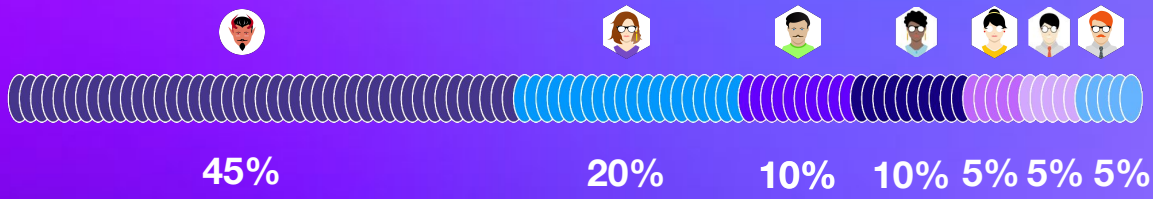
# Jury selection with secure proof-of-personhood




Jury:

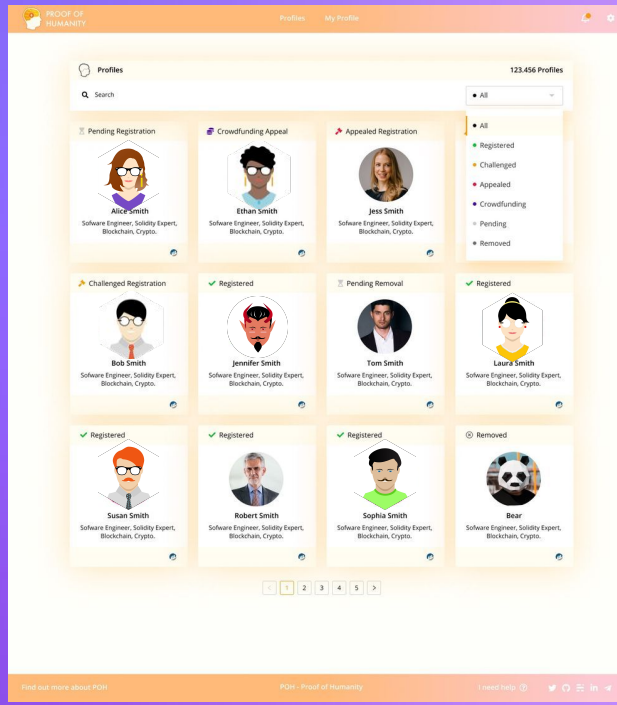


# Jury selection with secure proof-of-personhood

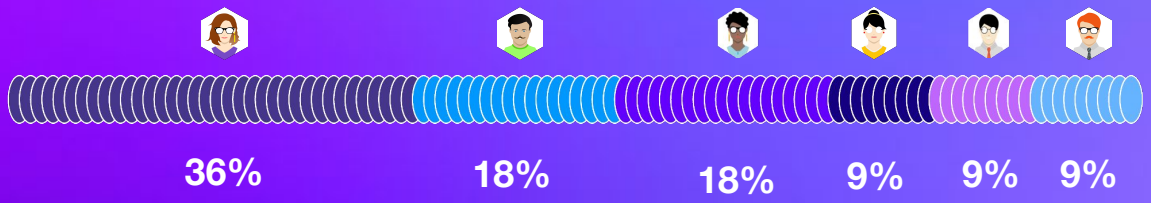



Jury:



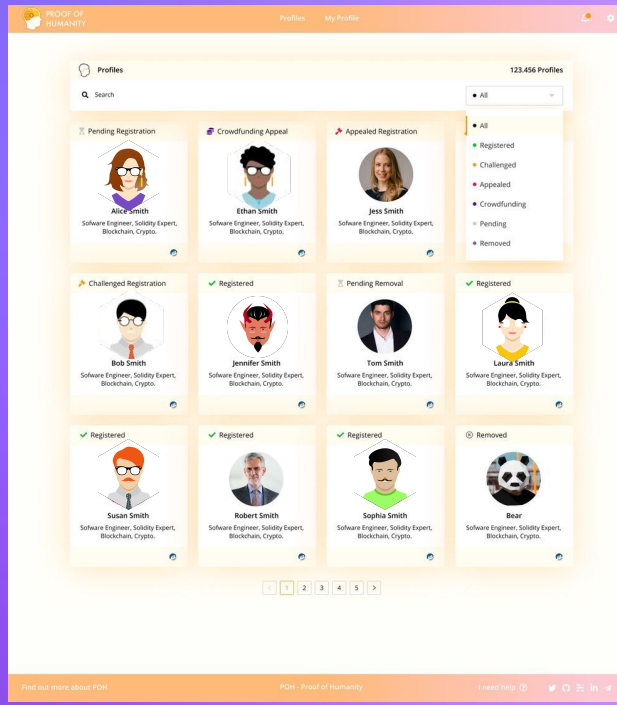


# Jury selection with secure proof-of-personhood

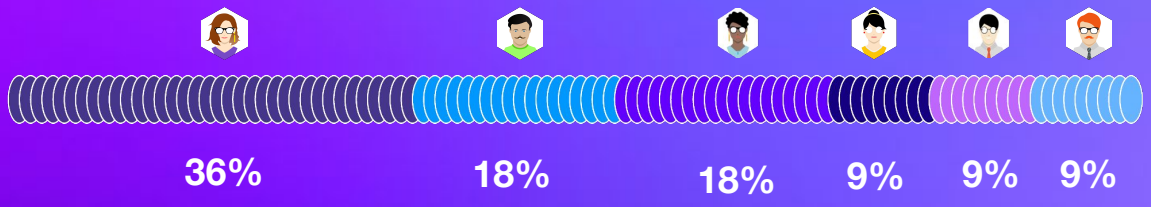



Jury:

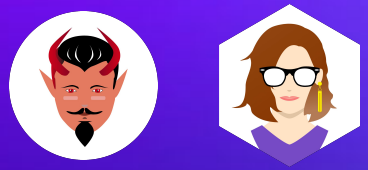


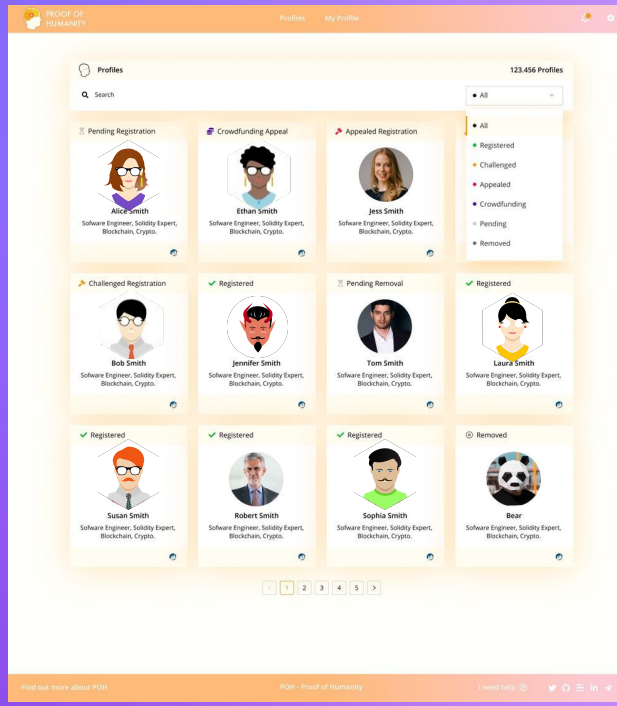


# Jury selection with secure proof-of-personhood

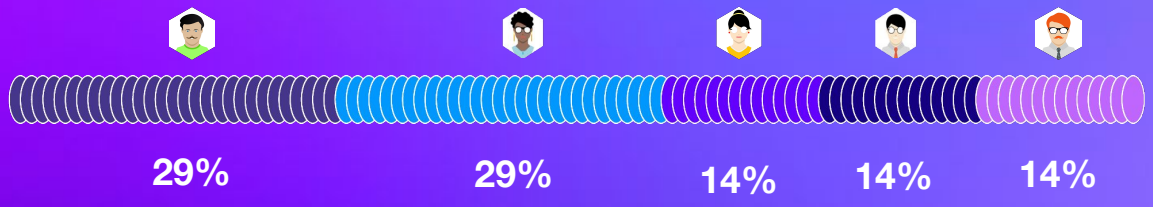



Jury:

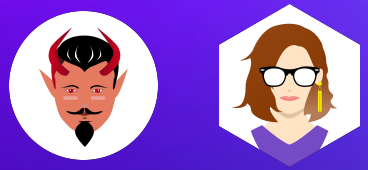


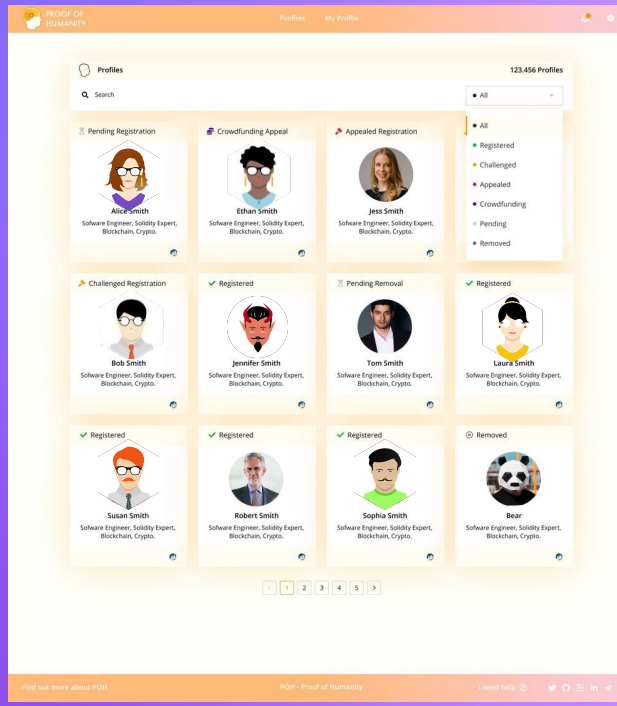


# Jury selection with secure proof-of-personhood

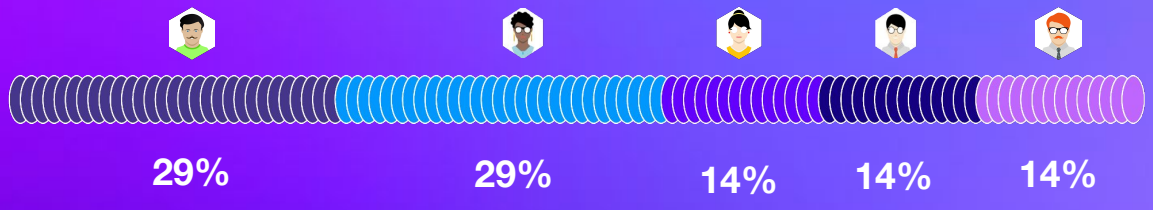



Jury:





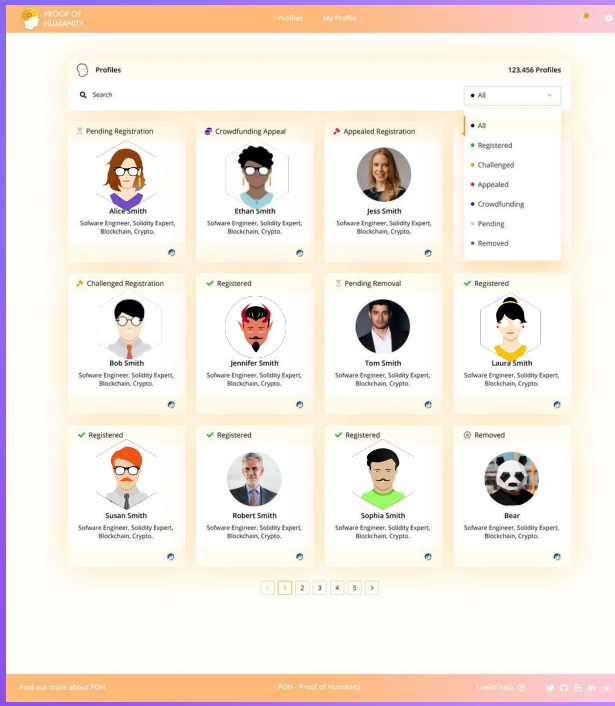
# Jury selection with secure proof-of-personhood



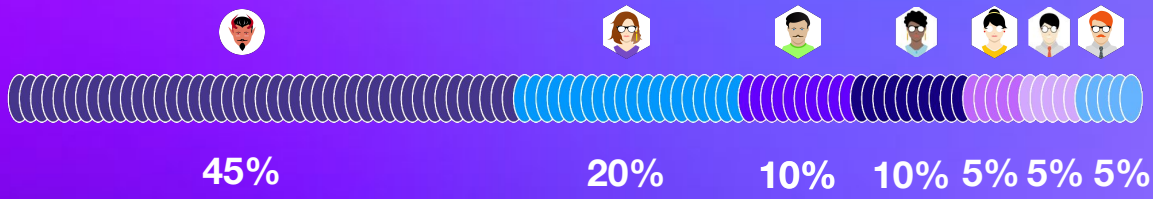

Jury:



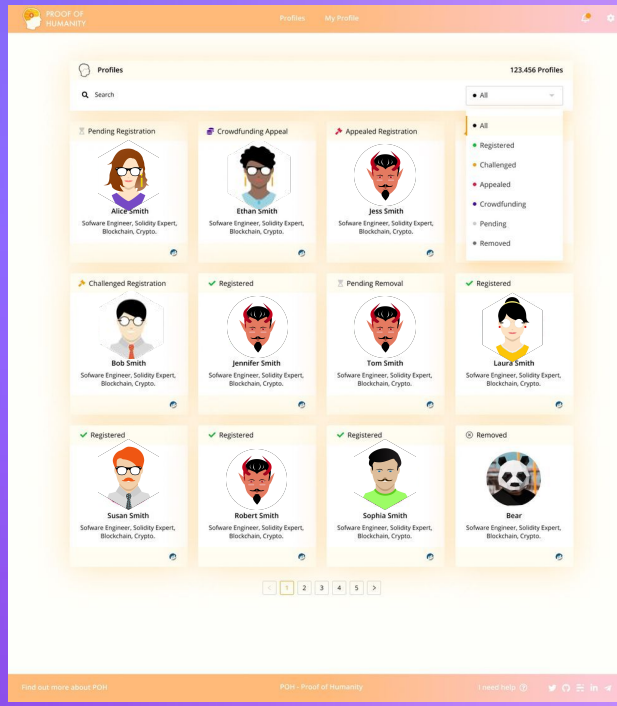
**0% overall chance of getting >= 2 ‘s**



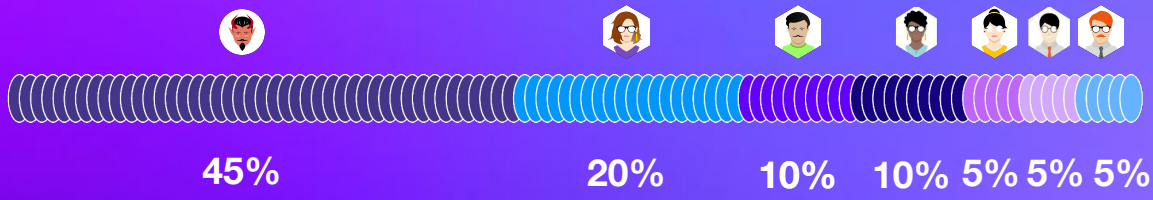
# Jury selection with insecure proof-of-personhood




Jury:

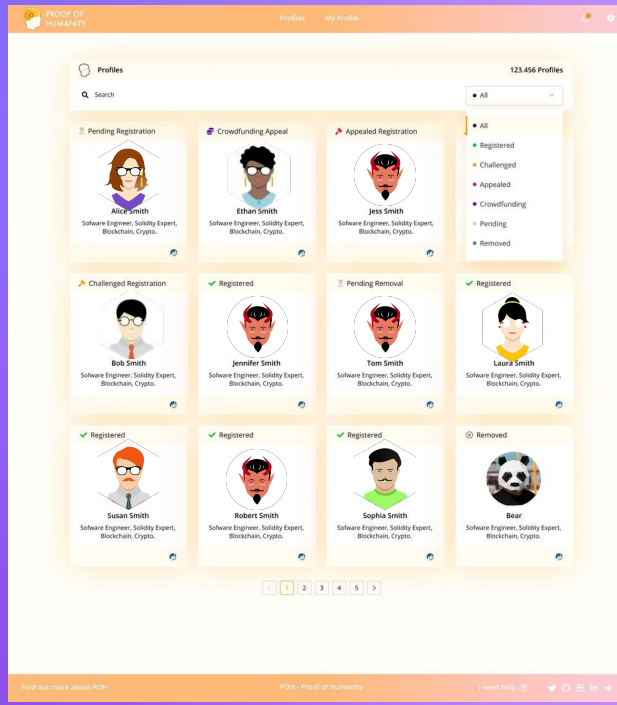


# Jury selection with insecure proof-of-personhood

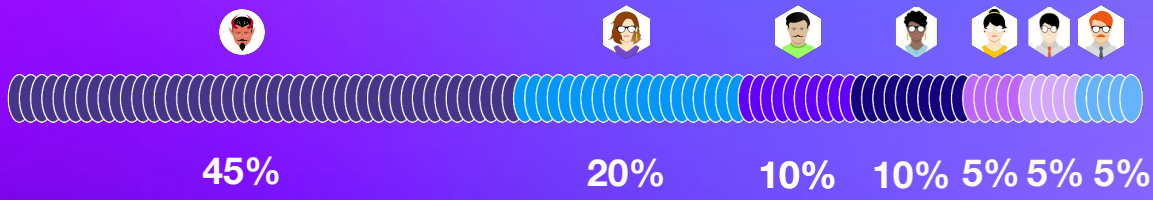



Jury:

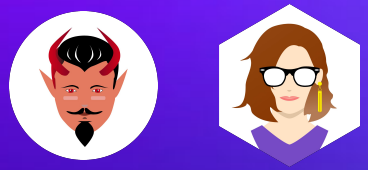


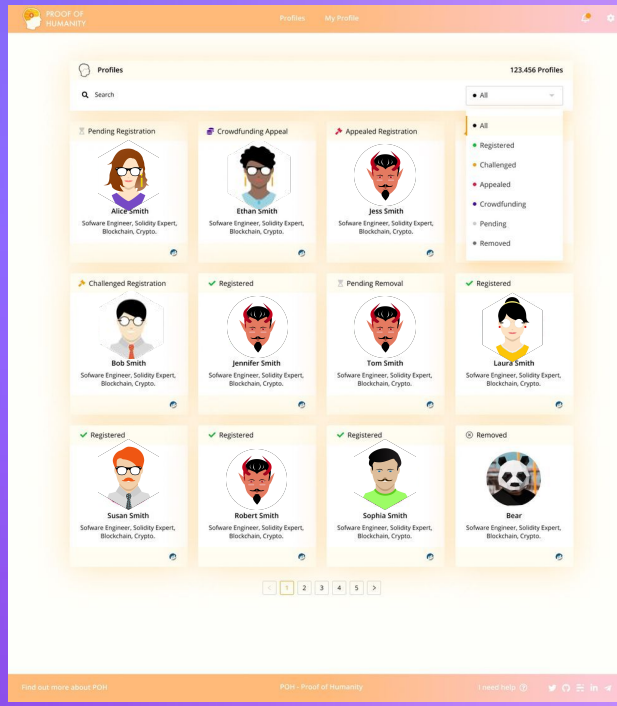


# Jury selection with insecure proof-of-personhood

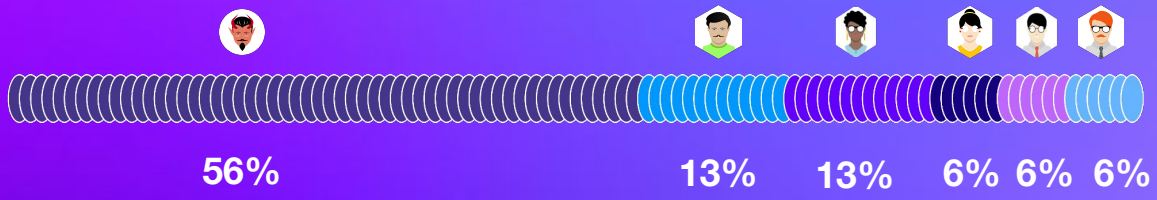



Jury:

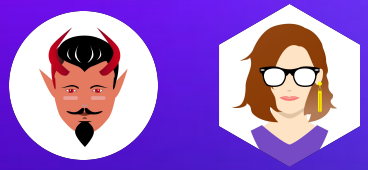


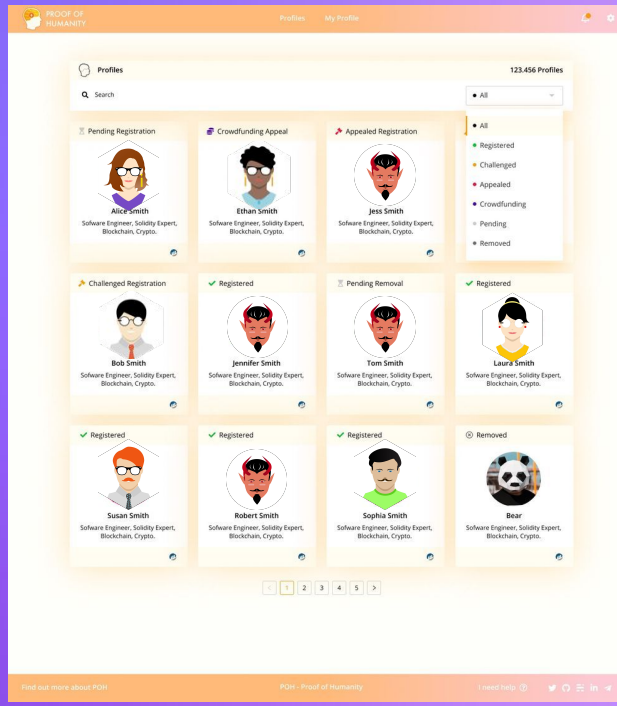


# Jury selection with insecure proof-of-personhood




Jury:





# Jury selection with insecure proof-of-personhood




Jury:



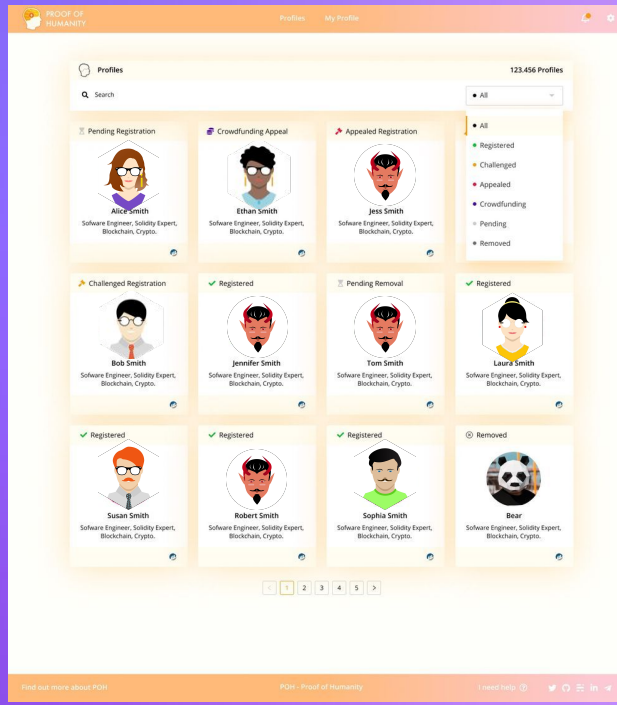


56% chance of another



44% chance of someone else

< 52% overall chance of getting >= 2  ‘s



# Jury selection with proof-of-personhood

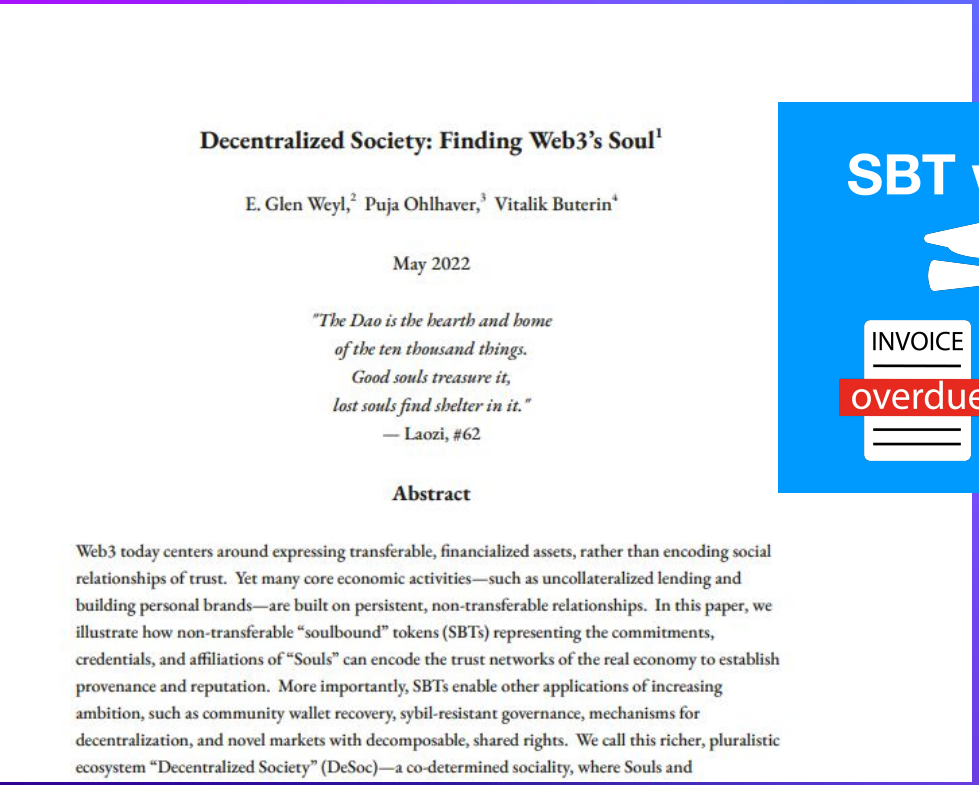
## SECURITY MODEL

**Proof-of-personhood protocol secure**

**OR**

**No participant controls more x% of stake***

*x depends on number of jury spots, distribution of other stakers



**Decentralized Society: Finding Web3's Soul**[1]

E. Glen Weyl,[2] Puja Ohlhaver,[3] Vitalik Buterin[4]

May 2022

*"The Dao is the hearth and home*
*of the ten thousand things.*
*Good souls treasure it,*
*lost souls find shelter in it."*
— Laozi, #62

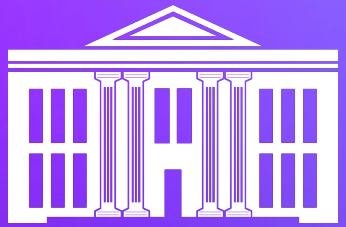
**Abstract**

Web3 today centers around expressing transferable, financialized assets, rather than encoding social relationships of trust. Yet many core economic activities—such as uncollateralized lending and building personal brands—are built on persistent, non-transferable relationships. In this paper, we illustrate how non-transferable "soulbound" tokens (SBTs) representing the commitments, credentials, and affiliations of "Souls" can encode the trust networks of the real economy to establish provenance and reputation. More importantly, SBTs enable other applications of increasing ambition, such as community wallet recovery, sybil-resistant governance, mechanisms for decentralization, and novel markets with decomposable, shared rights. We call this richer, pluralistic ecosystem "Decentralized Society" (DeSoc)—a co-determined sociality, where Souls and




- ➔ Non-transferable
- ➔ Have to present the bad with the good (short of scraping entire wallet - starting over)

University





“Alice is an expert in Subject X”

SBT Wallet



Alice

SBT

SBT Community



SBT

SBT Wallet



Alice

SBT

Pending submissions:



“I am an expert in Subject X”

Deposit:



SBT Wallet



Alice

SBT

Pending submissions:



“I am an expert in Subject X”

Deposit:





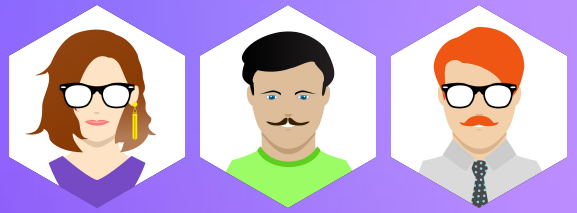


CHALLENGER

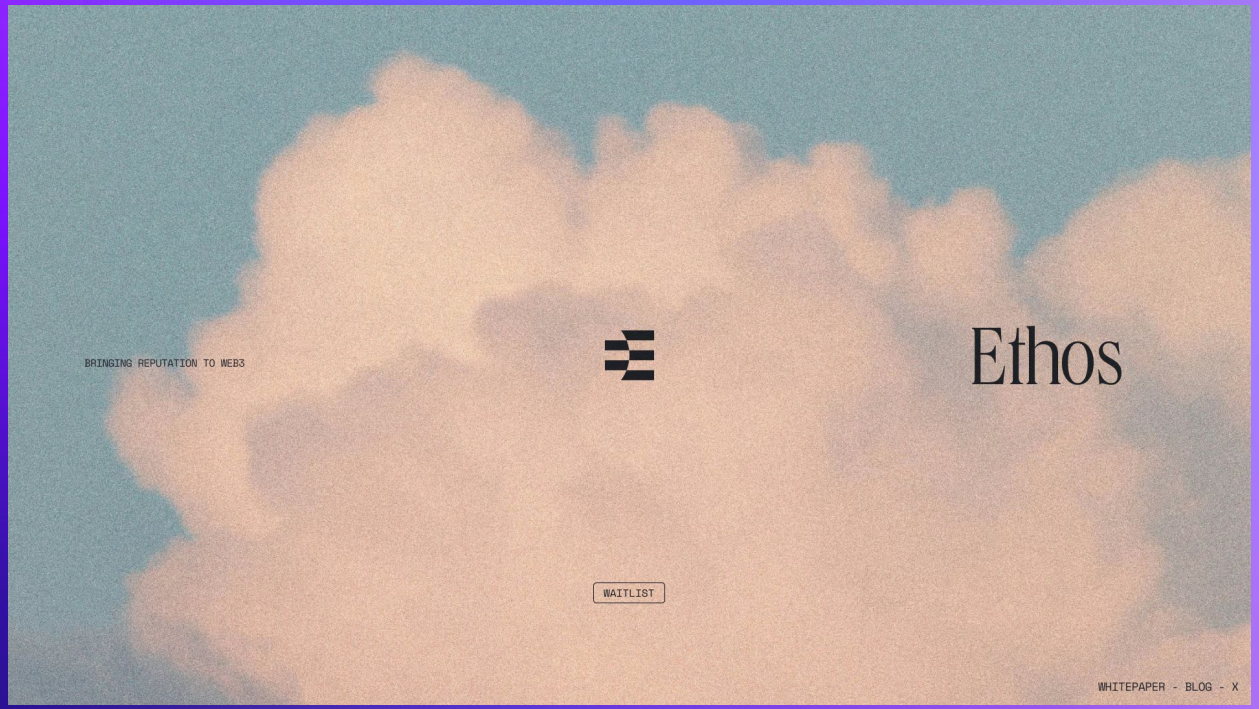
Deposit:



Jury



# Decentralized Credit Scores



# Agentic Reputation

Ethereum Improvement Proposals

All Core Networking Interface ERC Meta Informational

Draft Standards Track: ERC

## ERC-8004: Trustless Agents

### Discover agents and establish trust through reputation and validation


**Authors** Marco De Rossi (@MarcoMetaMask), Davide Crapis (@dcrapis) <davide@ethereum.org>, Jordan Ellis <jordanellis@google.com>, Erik Reppel <erik.reppel@coinbase.com>



**Created** 2025-08-13

**Discussion Link** https://ethereum-magicians.org/t/erc-8004-trustless-agents/25098

**Requires** EIP-155, EIP-712, EIP-721, EIP-1271


## Table of Contents


- Abstract
- Motivation
- Specification
    - Identity Registry
    - Reputation Registry
    - Validation Registry
- Rationale
- Security Considerations


Alice
SBT
Bob
SBT
Carl
SBT
Dave
SBT
Fred
PNK
PNK
PNK
PNK
PNK
Skills attested by SBT
SBT
Random
Selection

# Jury selection where required to have n of n SBTs

## SECURITY MODEL

Weakest SBT secure

OR

No participant controls more than near-majority x% of stake*

# Weighting by SBTs





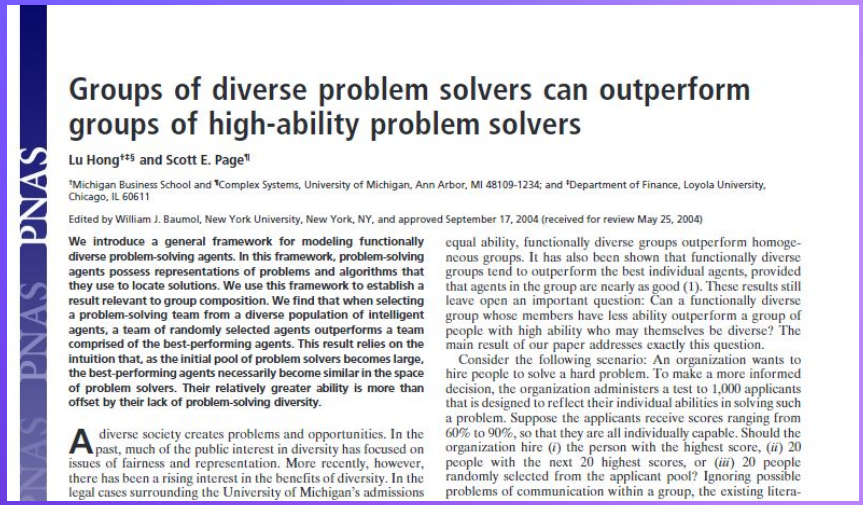







**POAP!**





PNAS PNAS PNAS

## Groups of diverse problem solvers can outperform groups of high-ability problem solvers

Lu Hong[‡§] and Scott E. Page[¶]

[‡]Michigan Business School and [¶]Complex Systems, University of Michigan, Ann Arbor, MI 48109-1234; and [§]Department of Finance, Loyola University, Chicago, IL 60611

Edited by William J. Baumol, New York University, New York, NY, and approved September 17, 2004 (received for review May 25, 2004)

**We introduce a general framework for modeling functionally diverse problem-solving agents. In this framework, problem-solving agents possess representations of problems and algorithms that they use to locate solutions. We use this framework to establish a result relevant to group composition. We find that when selecting a problem-solving team from a diverse population of intelligent agents, a team of randomly selected agents outperforms a team comprised of the best-performing agents. This result relies on the intuition that, as the initial pool of problem solvers becomes large, the best-performing agents necessarily become similar in the space of problem solvers. Their relatively greater ability is more than offset by their lack of problem-solving diversity.**

A diverse society creates problems and opportunities. In the past, much of the public interest in diversity has focused on issues of fairness and representation. More recently, however, there has been a rising interest in the benefits of diversity. In the legal cases surrounding the University of Michigan's admissions equal ability, functionally diverse groups outperform homogeneous groups. It has also been shown that functionally diverse groups tend to outperform the best individual agents, provided that agents in the group are nearly as good (1). These results still leave open an important question: Can a functionally diverse group whose members have less ability outperform a group of people with high ability who may themselves be diverse? The main result of our paper addresses exactly this question.

Consider the following scenario: An organization wants to hire people to solve a hard problem. To make a more informed decision, the organization administers a test to 1,000 applicants that is designed to reflect their individual abilities in solving such a problem. Suppose the applicants receive scores ranging from 60% to 90%, so that they are all individually capable. Should the organization hire (*i*) the person with the highest score, (*ii*) 20 people with the next 20 highest scores, or (*iii*) 20 people randomly selected from the applicant pool? Ignoring possible problems of communication within a group, the existing litera-

# Parallels to Quadratic Funding

Suppose participants $1,\ldots,K$ contribute $c_1^p,\ldots,c_K^p$ respectively to project p

Then project p receives a grant of :

$$\left(\sum_j \sqrt{c_j^p}\right)^2 - \sum_j c_j^p$$

# Parallels to Quadratic Funding

Proposal: Give  a grant

**Contribution**

 0

 1

 1

 4

 0

 1

 0

**Total Subsidy:**

$$(3 \cdot \sqrt{1} + \sqrt{4} + 3 \cdot \sqrt{0})^2 - 7 = 18$$

Proposal: Give  a grant

**Contribution**

 7

 0

 0

 0

 0

 0

 0

**Total Subsidy:**

$$(\sqrt{7} + 6 \cdot \sqrt{0})^2 - 7 = 0$$

# Weighting by SBTs in Quadratic Funding

## Idea:

- ➔ **Put users into groups based on SBTs**
- ➔ **Contributions from users from same group go under square root together (sort of like if same from same person)**

Beyond Collusion Resistance:
Leveraging Social Information for Plural Funding and Voting

Joel Miller* E. Glen Weyl† Leon Erichsen‡

December 2022

**Abstract**

In its 2018 introduction[2] and subsequent work[8], Quadratic Funding (QF) emerged as a uniquely optimal design for the democratic provision of public goods under the assumption of atomized participants with perfect rationality. In this paper, we aim to move past this rational and atomized portrayal of human beings and explore pluralistic QF mechanisms that recognize the influence of social connections among participants and incentivize cooperation across social differences. With this goal in mind, we define *collusion resistance* as a criterion against disproportional power accumulation in Quadratic Funding models that can arise from pre-existing participant relationships and argue that collusion resistant QF and plural QF are two sides of the same coin. Next, we evaluate various iterations of Quadratic Funding, testing their collusion resistance and other social and technical issues. Finally, we propose several new mechanisms, including Connection-Oriented Cluster Match, which satisfies our definition of collusion resistance. Our findings show great potential for making Quadratic Funding more pluralistic. They might also guide principles and practices of computational design that bridge the epistemological divides between classical economics and social reality.

# Weighting by SBTs in Quadratic Funding

**Total amount of funding for project =**

$$\left( \sum_{\text{groups g}} \sqrt{\sum_{\text{participant } j \in g} c_j} \right)^2$$

# Weighting by SBTs in QF

| | Contribution | Group |
|---|---|---|
|  | 5 | 1 |
|  | 7 | 2 |
|  | 1 | 2 |
|  | 2 | 1 |

**Funding =**

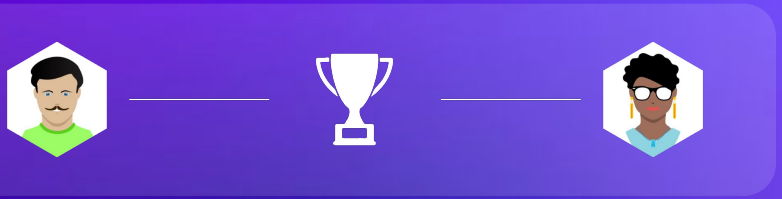
$$\left( \sum_{\text{groups } g} \sqrt{\sum_{\text{participant } j \in g} c_j} \right)^2$$

$$= (\sqrt{5+2} + \sqrt{7+1})^2$$

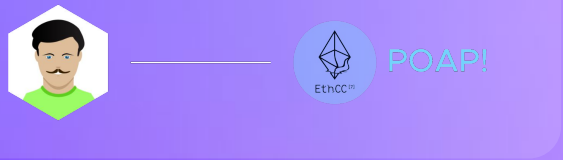
# Weighting by SBTs in Kleros



2 5



7/2 1




7/2

# Weighting by SBTs in Kleros

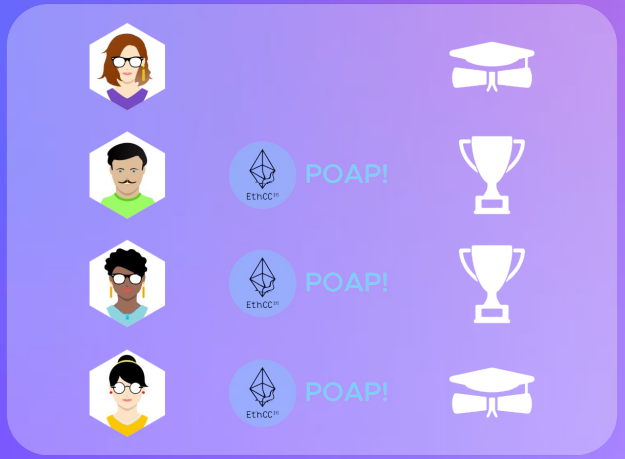
**Total (re-weighted) stake** $\sqrt{5+2}+\sqrt{7/2+1}+\sqrt{7/2}=6.64$

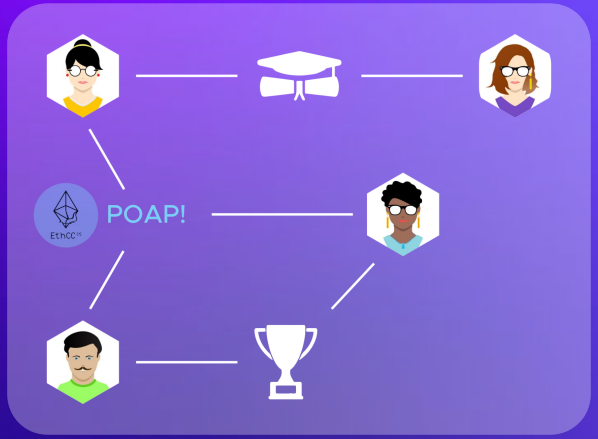
| | (Re-weighted) stake | % chance of getting drawn with SBT weighting |
|---|---|---|
|  | $\frac{5}{\sqrt{5+2}}=1.89$ | 1.89/6.64 |
|  | $\frac{7/2}{\sqrt{7/2}}+\frac{7/2}{\sqrt{7/2+1}}=3.52$ | 3.52/6.64 |
|  | $\frac{1}{\sqrt{7/2+1}}=.47$ | .47/6.64 |
|  | $\frac{2}{\sqrt{5+2}}=.76$ | .76/6.64 |

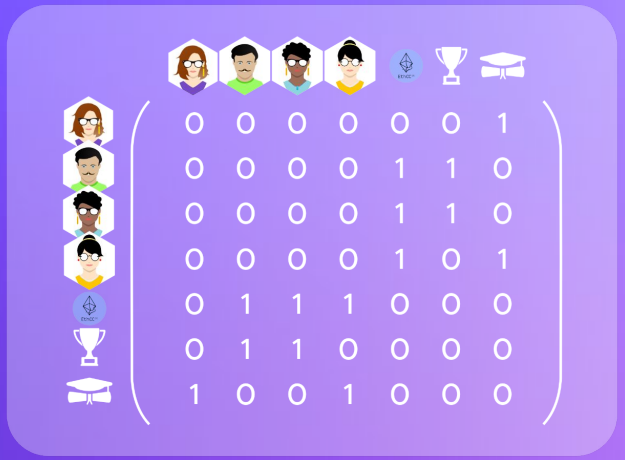
# Kleros with SBT weighting – **Security Model**

- Secure unless attacker has both large proportion of Kleros staking tokens AND the right combination of SBTs
- If mechanism to distribute SBTs secure, generally harder to attack than 1.0 mechanism
- If mechanism to distribute SBTs is broken, reduces staking attack threshold

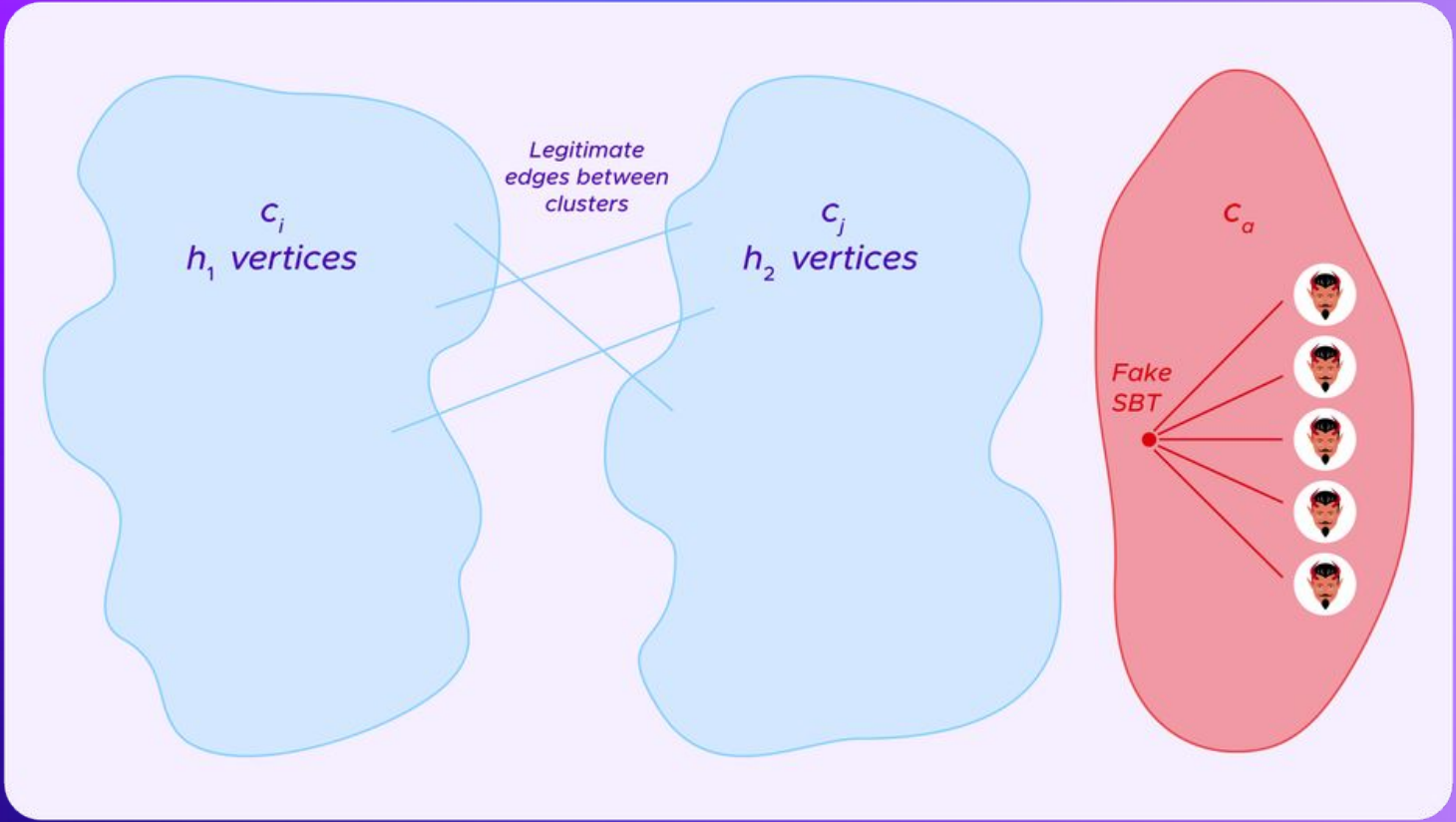
# Group using modern clustering algorithm: **spectral clustering**



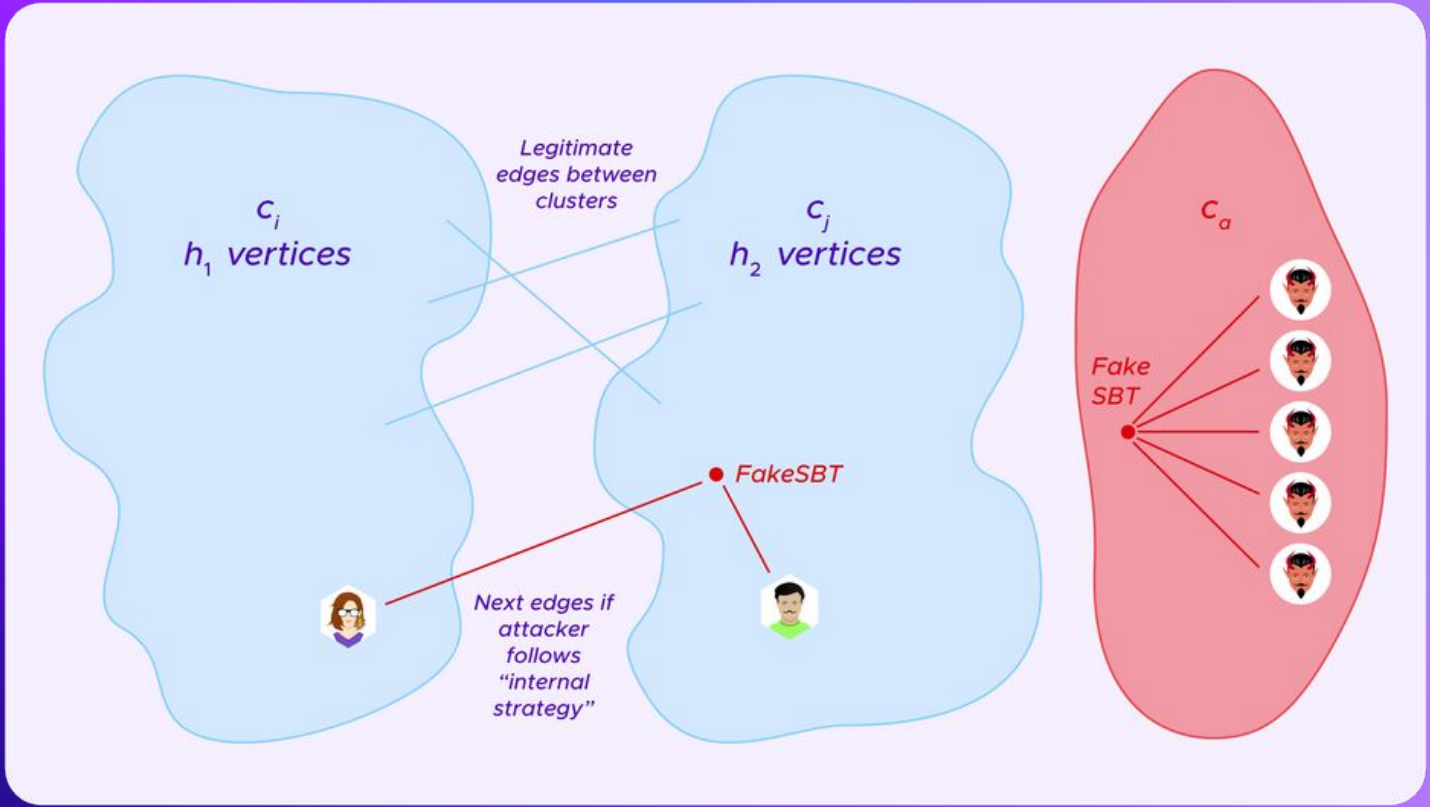



# Attacks on Spectral Clustering Weighting



# Attacks on Spectral Clustering Weighting



# Conclusions

- Lack of identity tools forces to weight by capital so not manipulable by Sybil attack
  - Leads to hyper-financialization
- New blockchain identity tools such as Proof-of-Personhood and Soulbound Tokens offer tools to build systems with different security models
- Ideally want economic and social defenses to be complementary

KLEROS

kleros.io

25 September 2026

# The Role of Identity

## in Decentralized Dispute Resolution Protocols

**William George**
Research Lead @ Kleros

@kleros_io

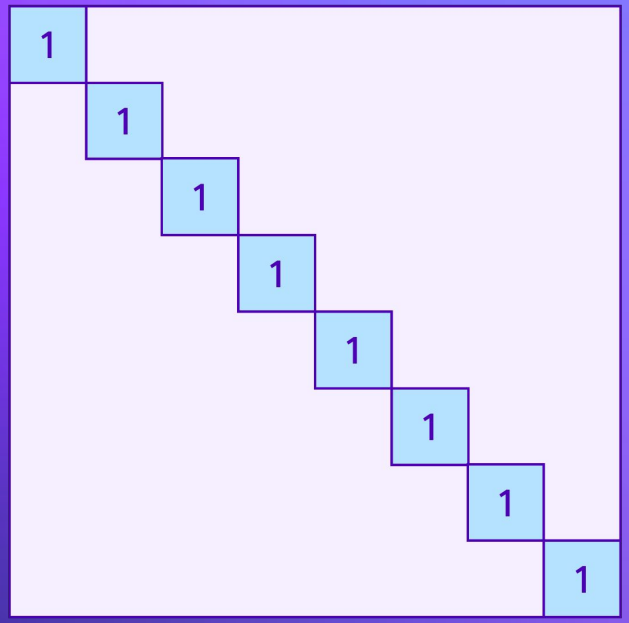
HOSTED BY THOMSON REUTERS INCUBATOR

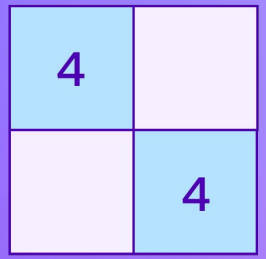
European Commission | Winner of the Blockchains for Social Good Prize from the European Union's Horizon 2020 research and innovation programme

bpifrance

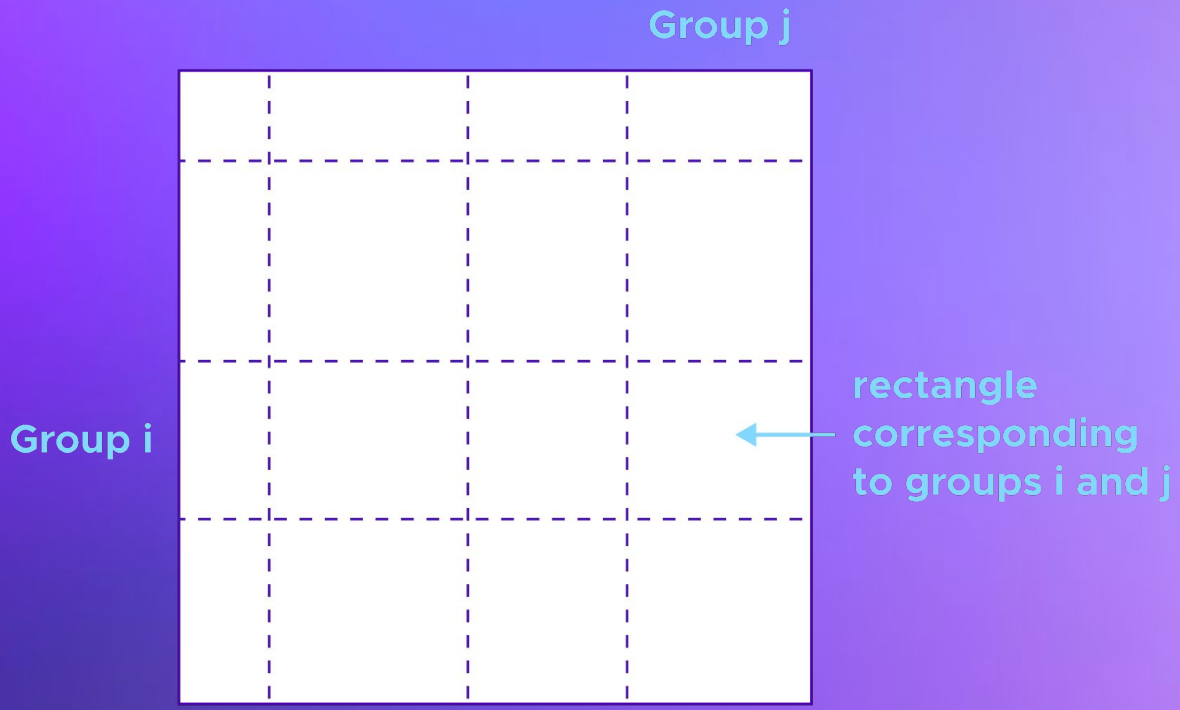
BLOCKCHAINGOV

# Parallels to Quadratic Funding

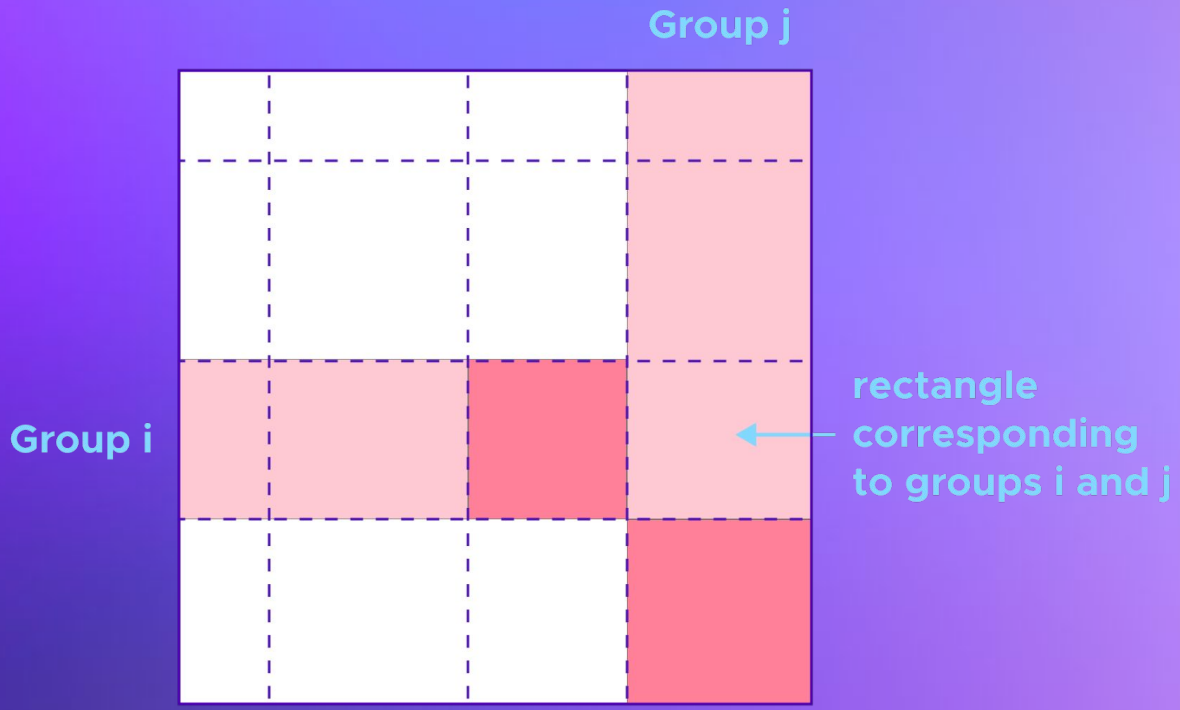


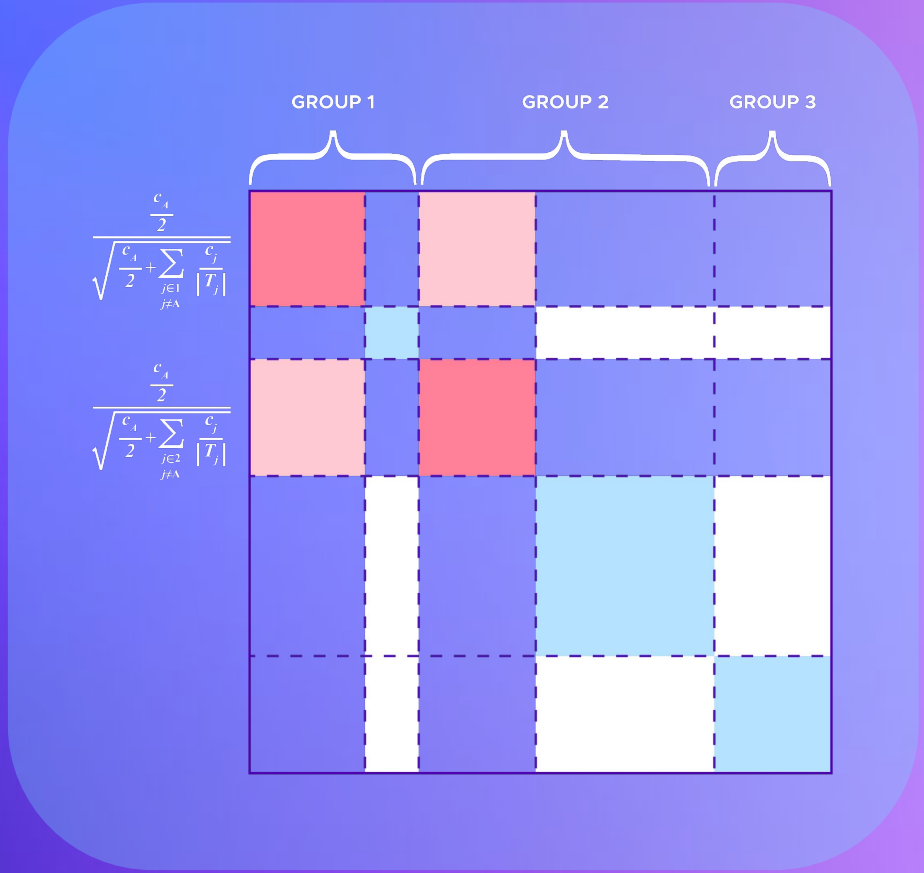
# Weighting by SBTs in QF

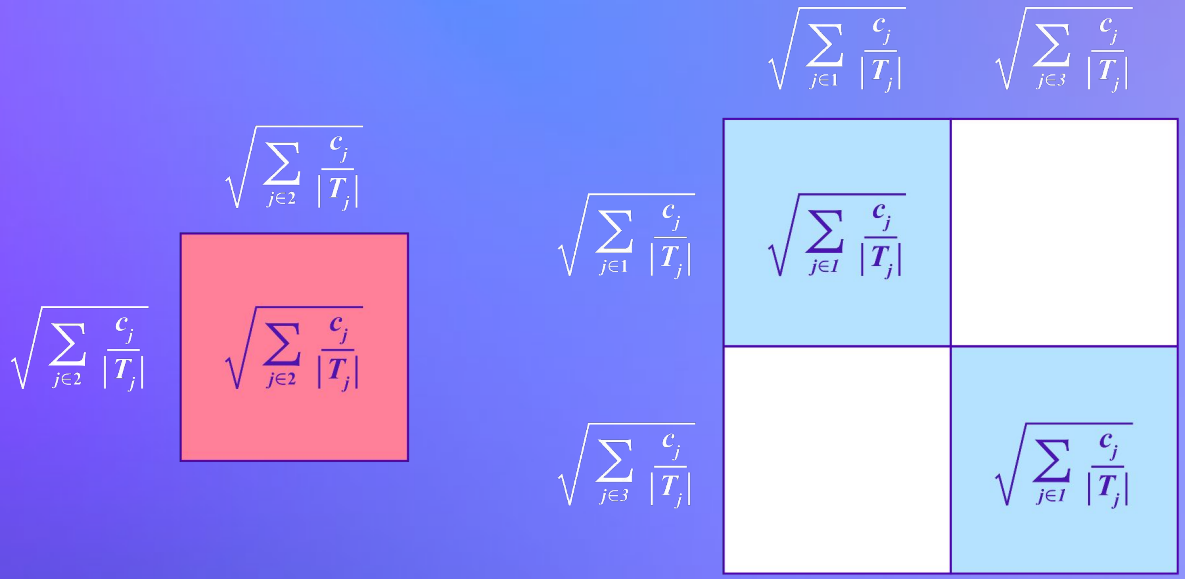


# Weighting by SBTs in QF



# Full square of stakes decomposes into “honest square” and “attacker square”






**Attacker has > 50% chance of being selected in given draw if and only if attacker's square larger than honest square**