Economic Incentives and Souls in Schelling-point Based Oracles

WILLIAM GEORGE
Research Lead, Kleros

kleros.io

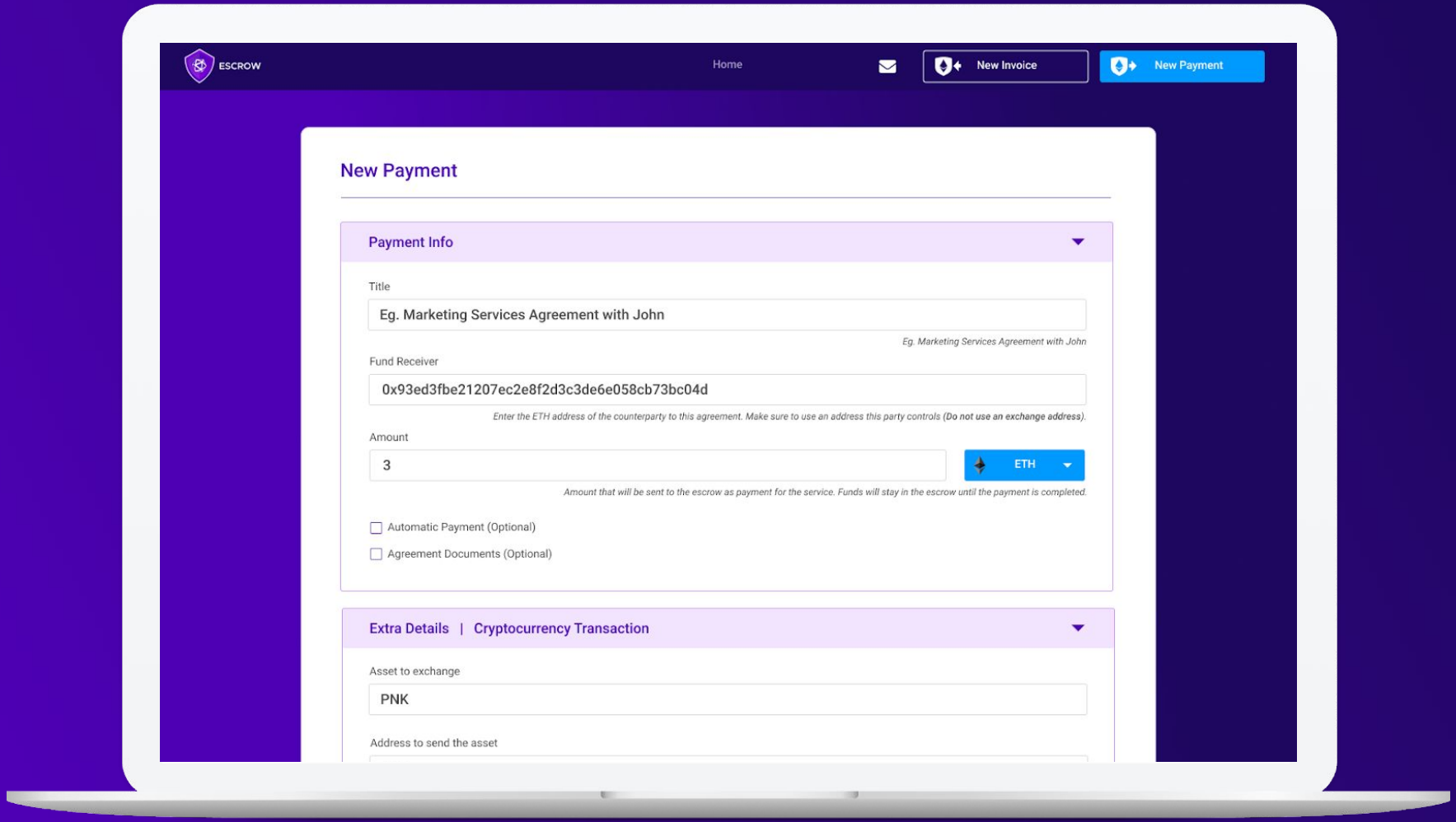
KLEROS

HOSTED BY
THOMSON REUTERS
INCUBATOR

European Commission

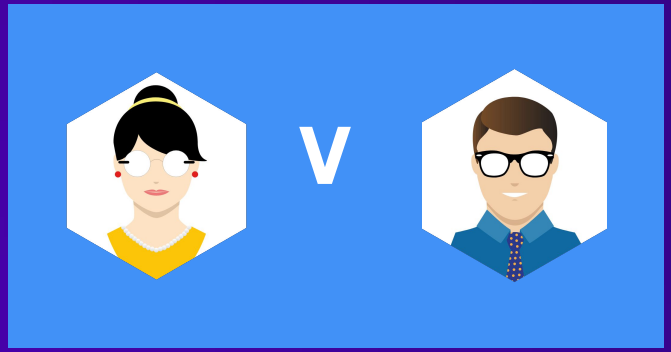
Winner of the Blockchains for Social Good Prize from the European Union's Horizon 2020 research and innovation programme

bpifrance


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

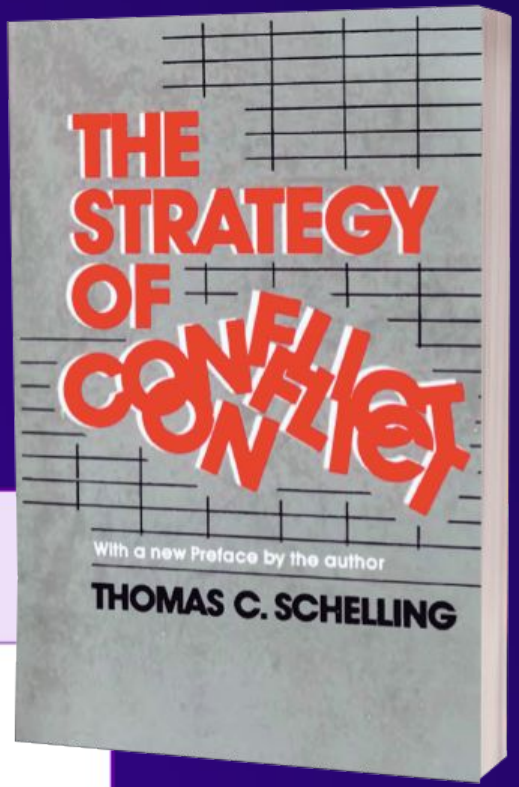

KLEROS


KLEROS


V

Schelling Point

- “solution that people tend to use in the absence of communication because it seems natural, special, or relevant to them”



The majority vote \ You vote	X	Y
X	1	-1
Y	-1	1

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

Decentralized Society: Finding Web3's Soul[1]

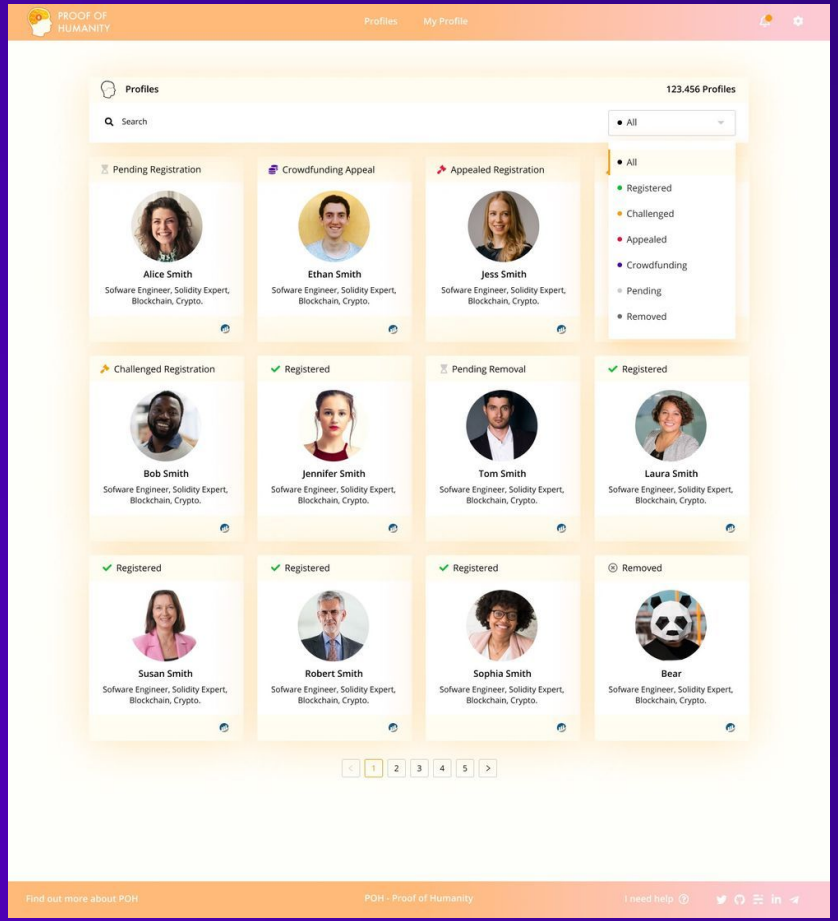
E. Glen Weyl,[2] Puja Ohlhaver,[3] Vitalik Buterin[4]

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

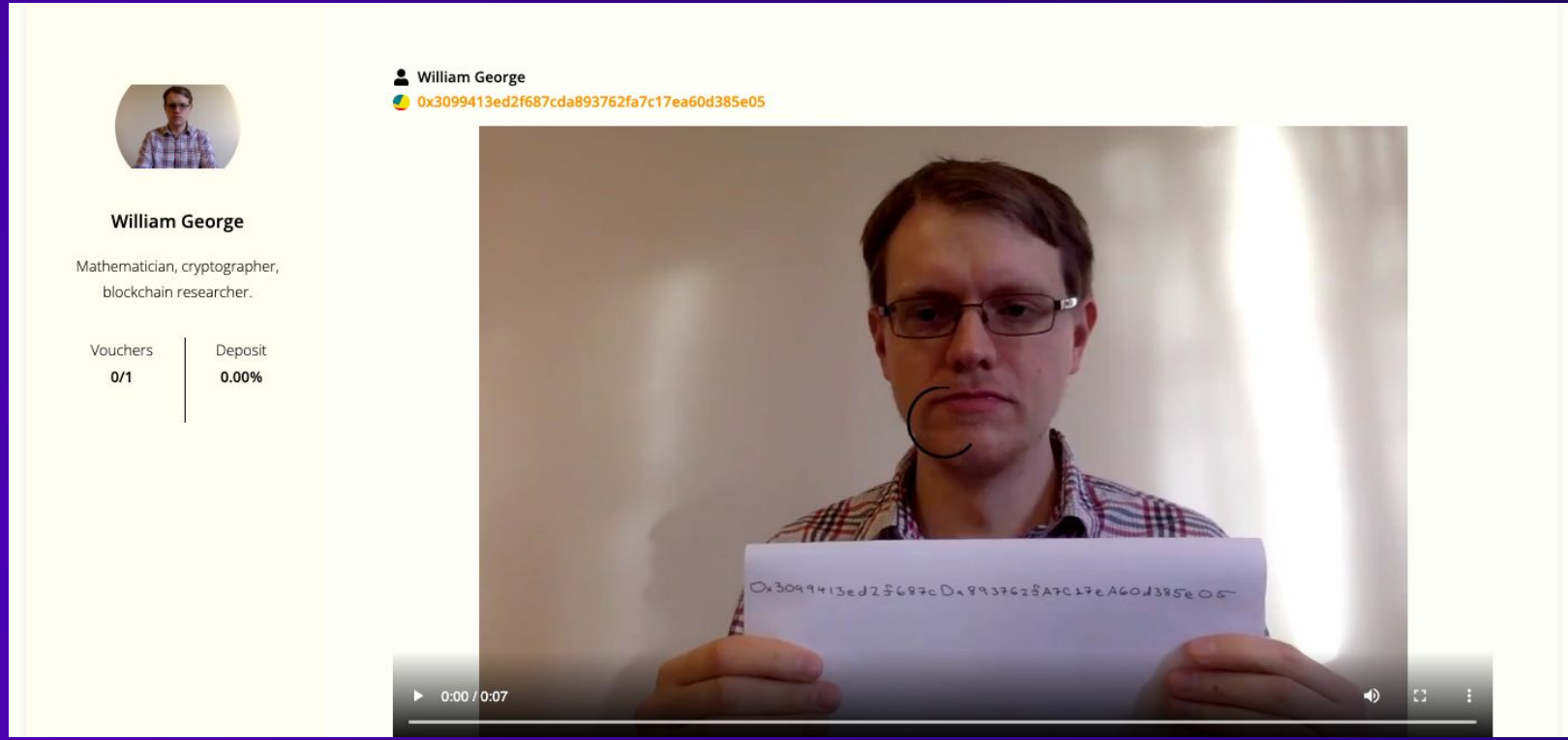
Abstract

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



PROOF OF HUMANITY

- → Proof of personhood/Sybil resistance tool
- → Curated list, if dispute over eligibility -> Kleros dispute

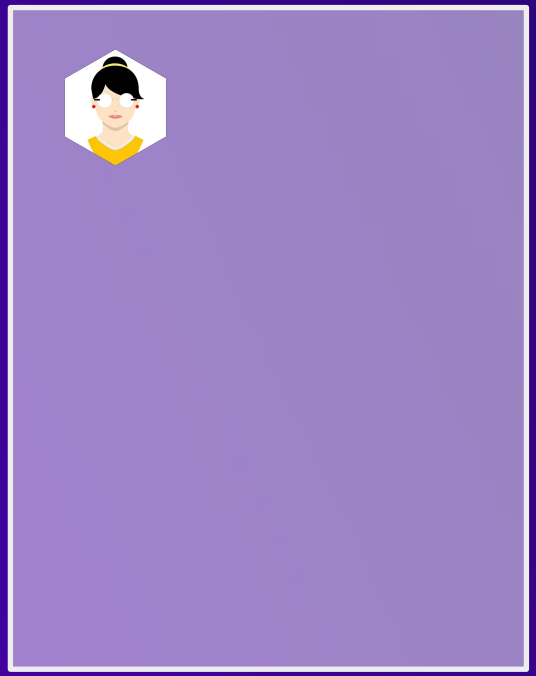


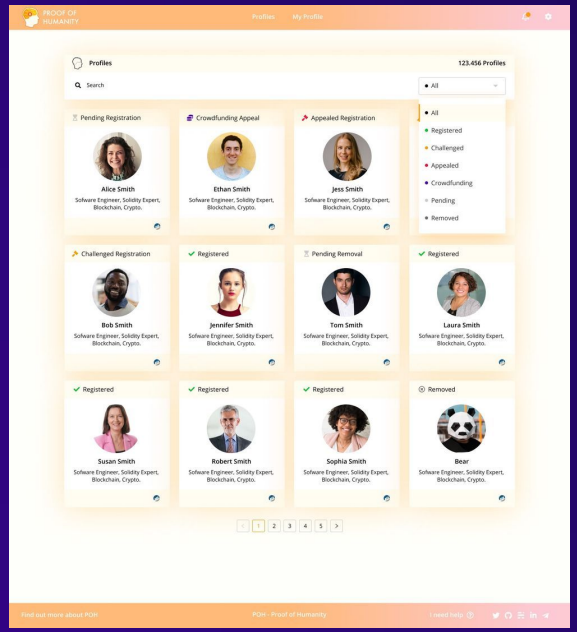
“I certify that I am a real human, and that I am not already registered in this registry.’”

Pending submissions:

“I am a human not already on the registry”

Deposit:





Pending submissions:

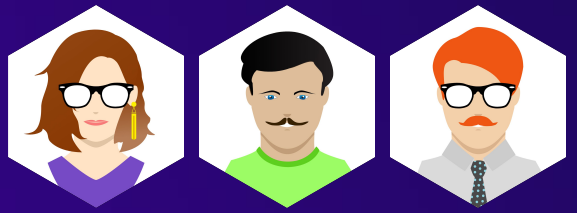
“I am a human not already on the registry”

Deposit:

CHALLENGER

Deposit:

Jury



Removal Request

A request to remove a submission in "Registered" state from the list can be made at any time by anyone submitting a deposit.

The removal requester has to either:

- Provide evidence that the above-cited acceptance criteria are not fulfilled by the submission.
  - Example: Send the following removal request
    Evidence name: This user video is a deep-fake
    Evidence description: You will find in the attached file an analysis report proving that this video is deep-faked.

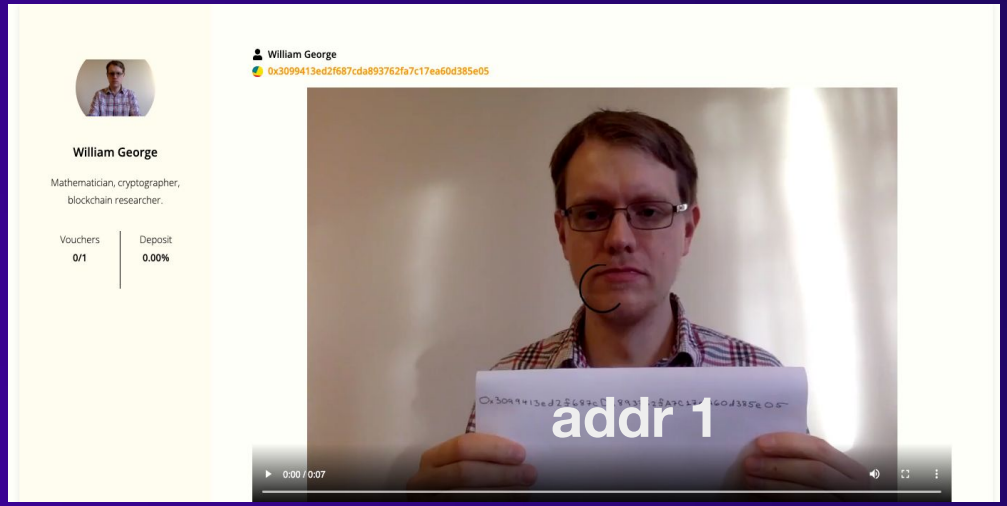
- Or provide evidence that he is the submitter and wants to voluntarily remove his submission.
  - Example 1: Send a removal request from the same address as the submitter.
    - *Evidence Name: Self-removal of submission*
    - *Evidence Description: I am the submitter as proven by my address and I want to remove this submission.*

  - Example 2: Send a removal request from a different address than the submitter.
    - *Evidence Name: Self-removal of submission*
    - *Evidence Description: I am the submitter and I want to remove this submission. The video attached is a recording of myself saying the sentence "I want to remove my own submission from the Proof of Humanity registry."*

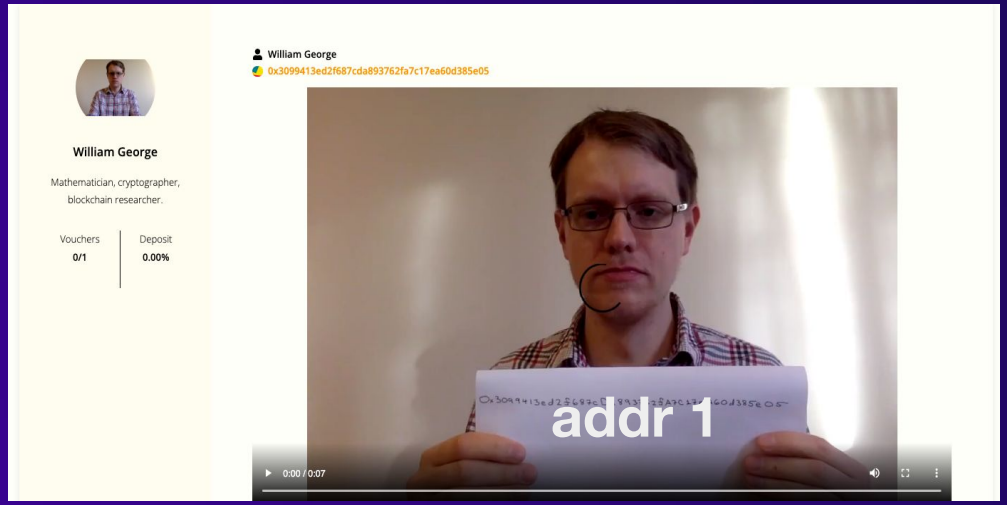
"Video attached is a recording of myself saying the sentence 'I want to remove my own submission from the Proof of Humanity registry.'"

SBT wallet

addr 1



SBT wallet



addr 1

addr 2

I want to remove my own submission. Resubmit with addr2.

“ is an expert in Subject X”

‘s SBT wallet

in X

Pending submissions:

“I am an expert in Subject X”

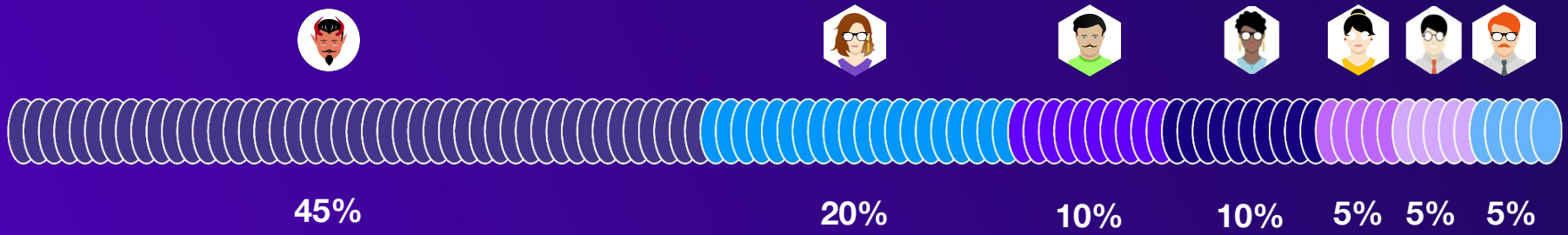
Deposit:

‘s SBT wallet

in X

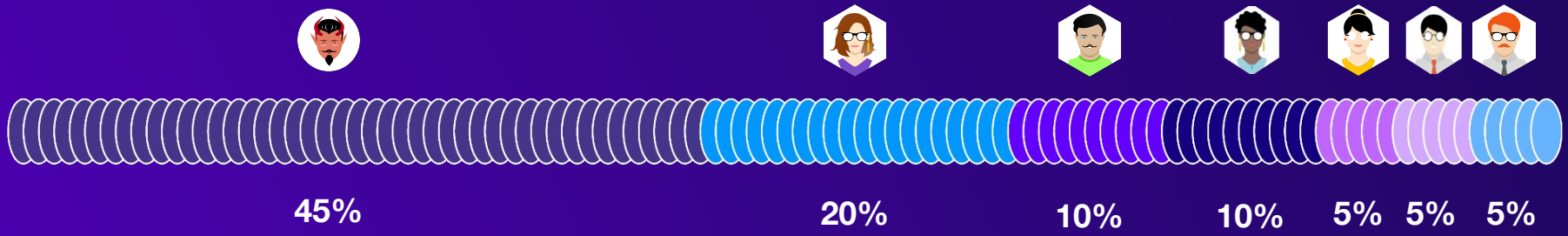
Effects of using proof-of-personhood and SBTs on attack resistance

Jury selection without proof-of-personhood



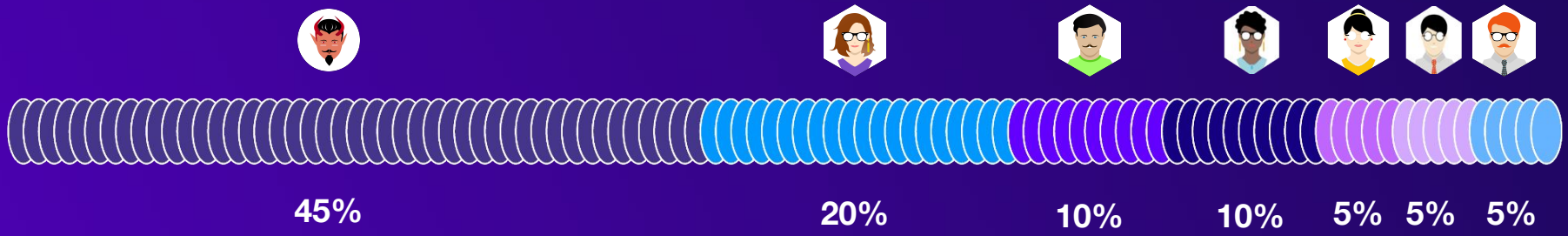
Jury:

Jury selection without proof-of-personhood



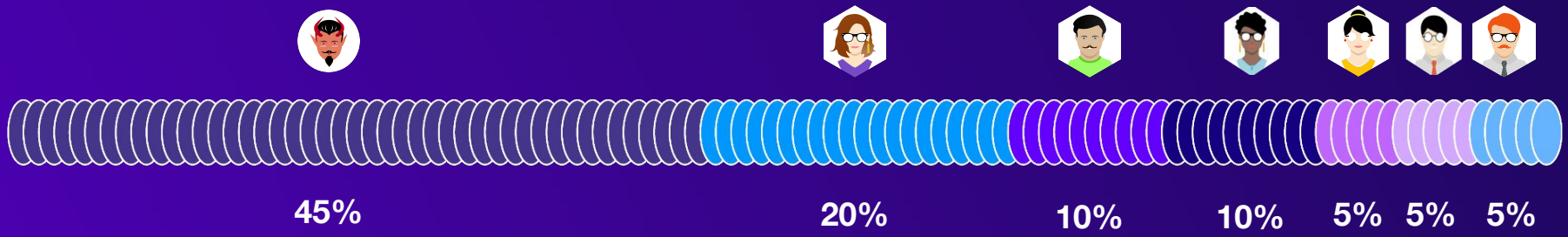
Jury:

Jury selection without proof-of-personhood



Jury:

Jury selection without proof-of-personhood



Jury:

45% chance of another

55% chance of someone else

42.5% overall chance of getting >= 2 ‘s

Jury selection without proof-of-personhood

Security Model

No participant controls more than 50% of stake

Quadratic Funding

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

Then project p receives a grant of :

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

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:

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

Quadratic Funding with secure proof-of-personhood

Proposal: Give  **a grant**

Contribution

 1

 0

 0

 0

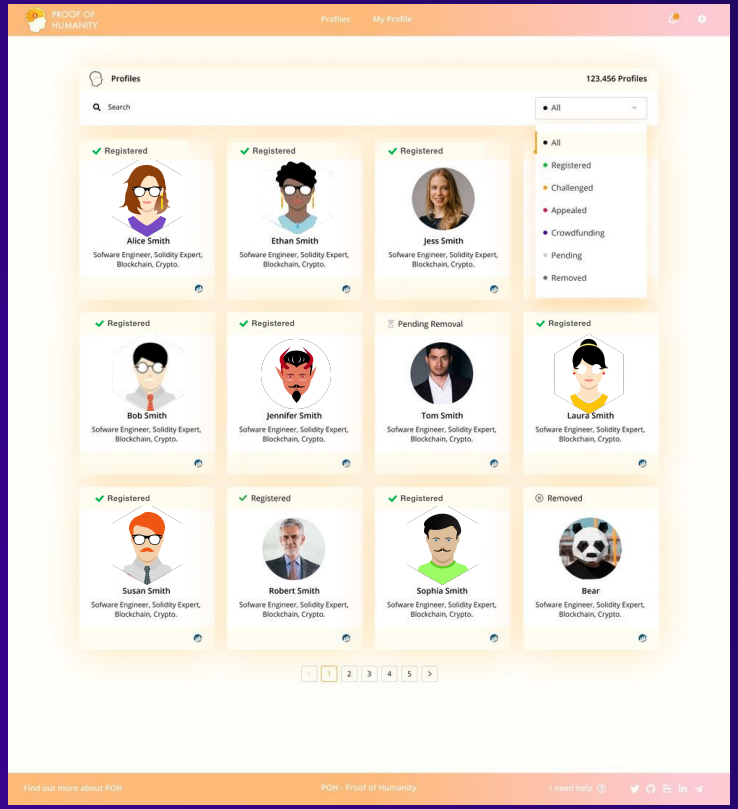
 0

 0

 0

Total Subsidy:

$$(\sqrt{1} + 6 \cdot \sqrt{0})^2 - 1 = 0$$



Quadratic Funding without (secure) proof-of-personhood

Proposal: Give  **a grant**

Contribution

1/N with each of N profiles

0

0

0

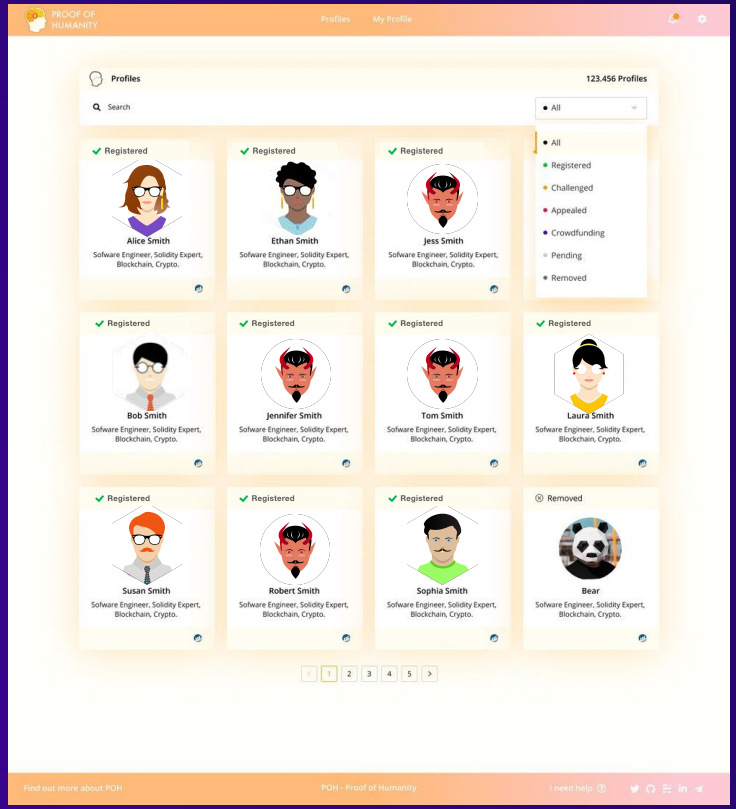
0

0

0

Total Subsidy:

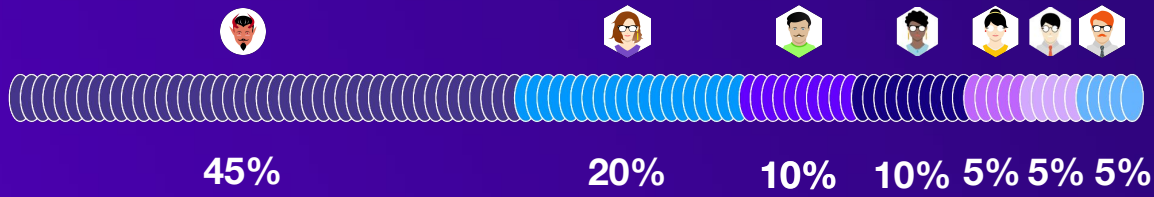
$$\left(N \cdot \sqrt{\frac{1}{N}} + 6 \cdot \sqrt{0}\right)^2 - 1 = N - 1$$



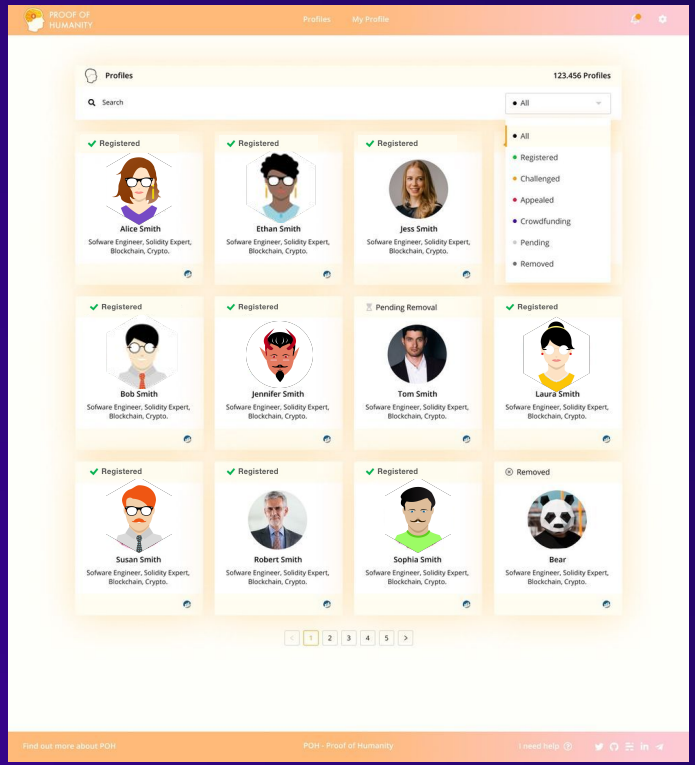
Quadratic Funding

Proof-of-personhood protocol secure

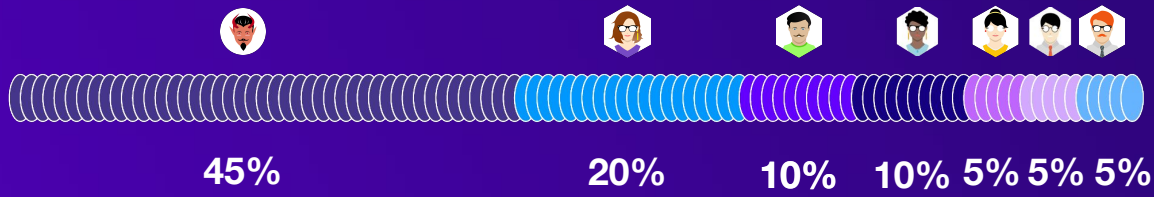
Jury selection with secure proof-of-personhood



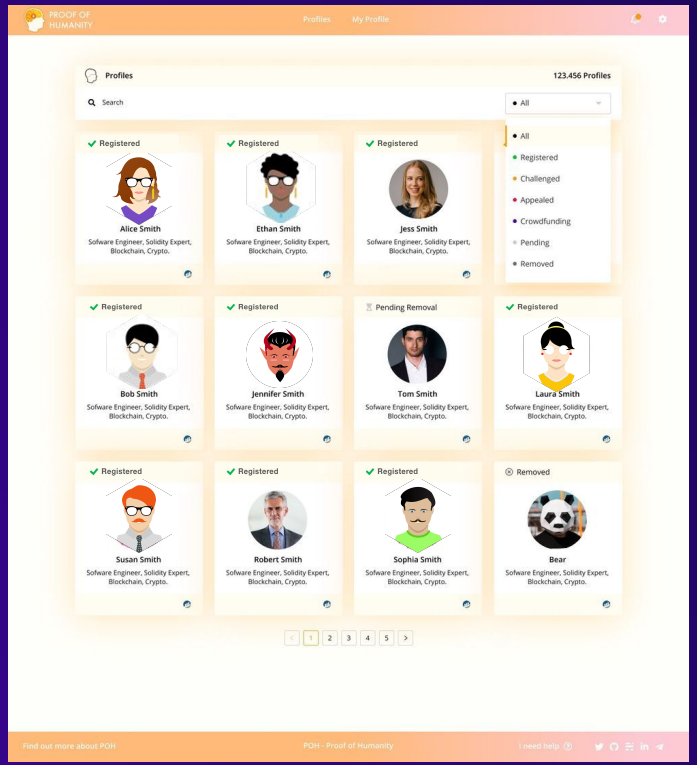
Jury:



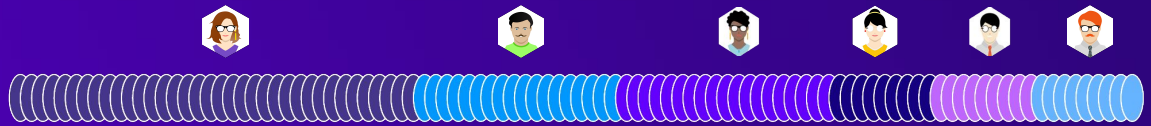
Jury selection with secure proof-of-personhood



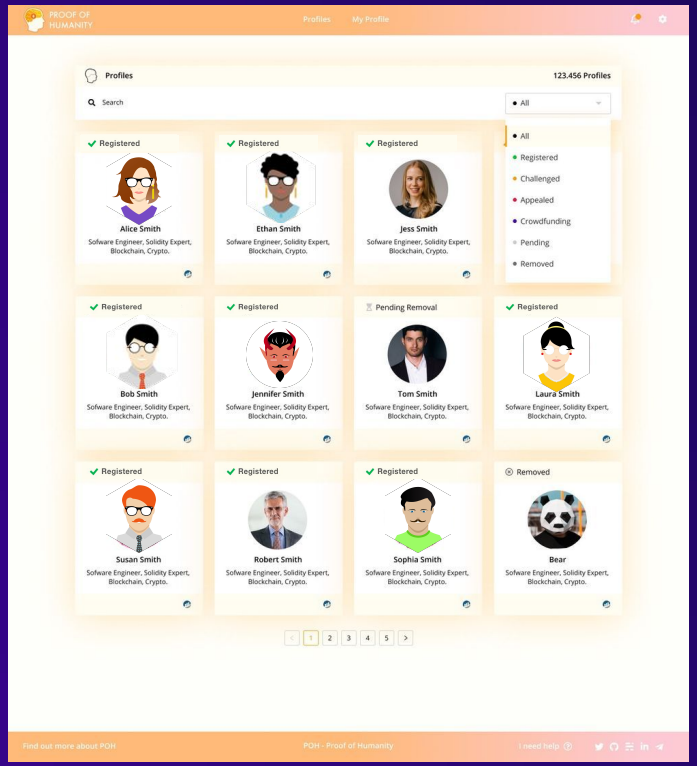
Jury:



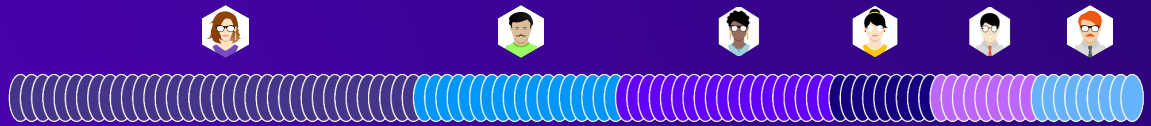
Jury selection with secure proof-of-personhood



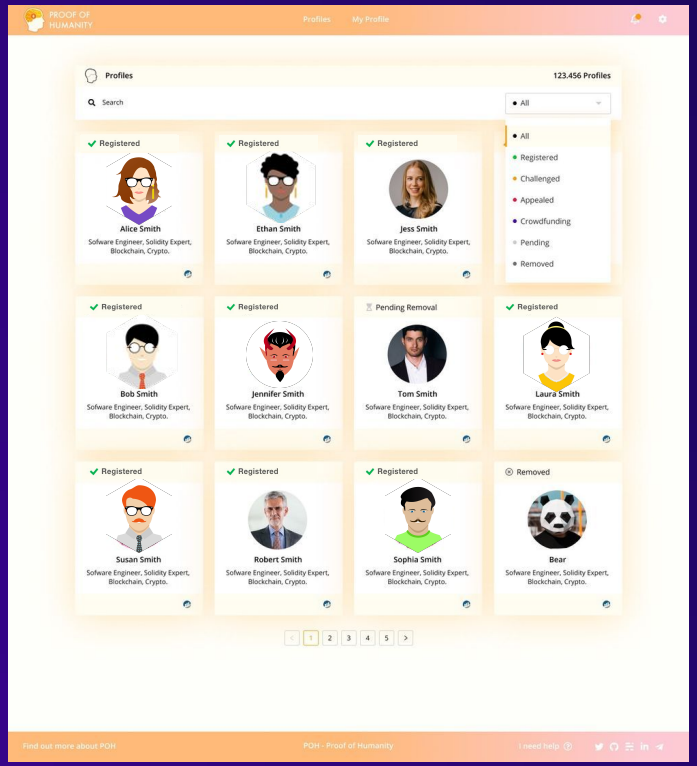
Jury:



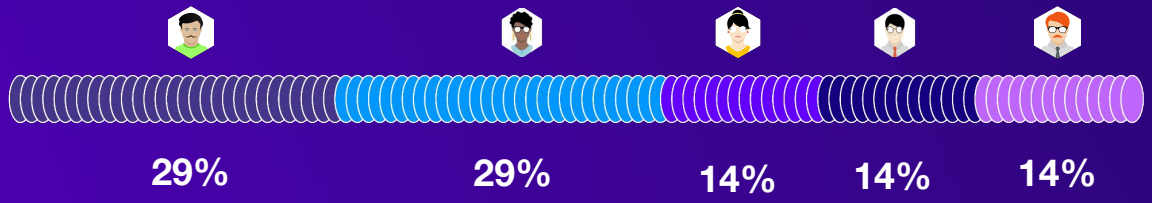
Jury selection with secure proof-of-personhood



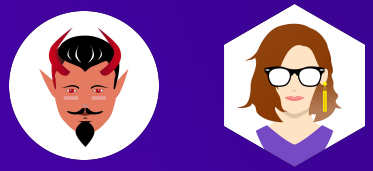
Jury:

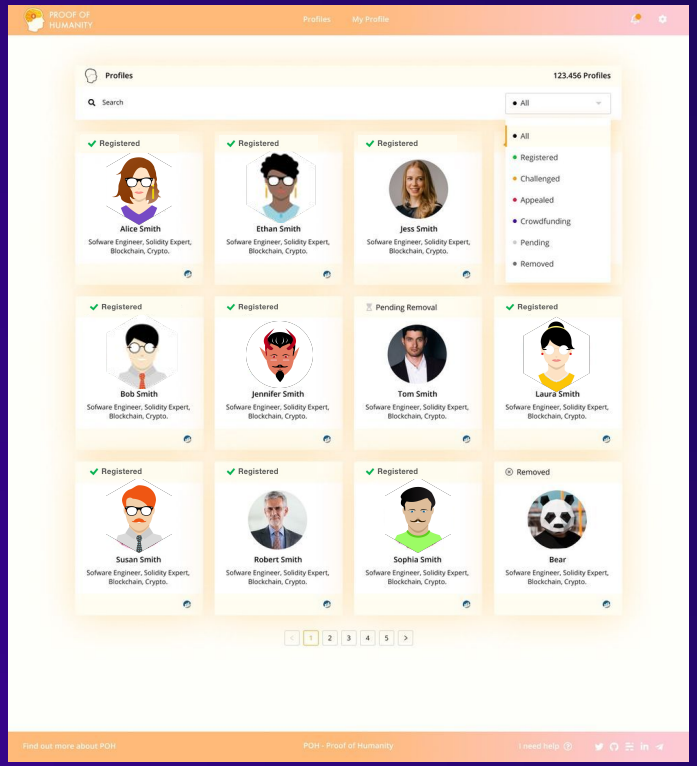


Jury selection with secure proof-of-personhood

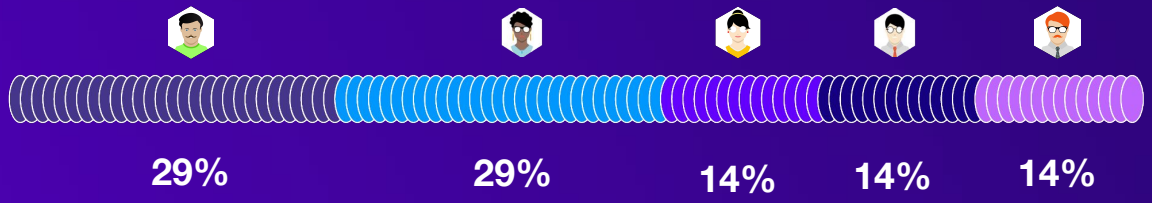


Jury:



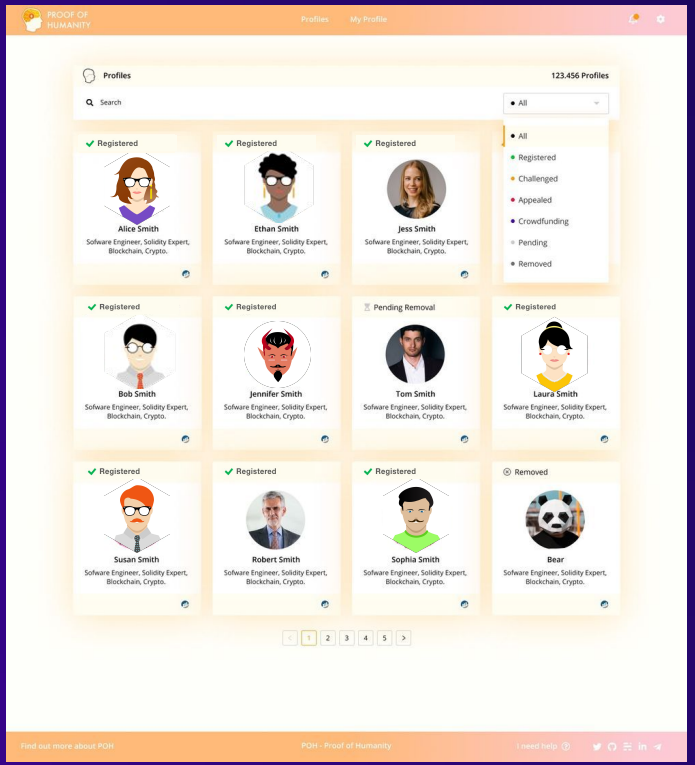


Jury selection with secure proof-of-personhood

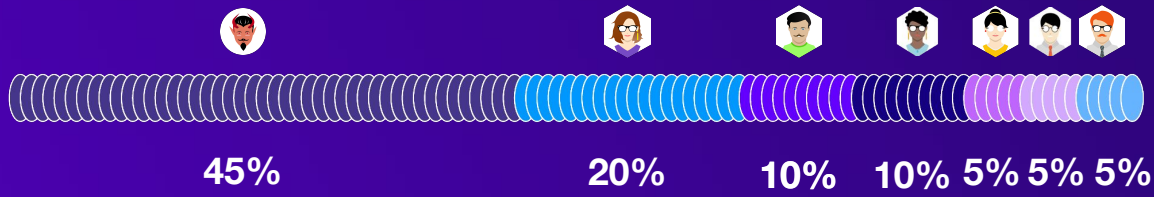


Jury:

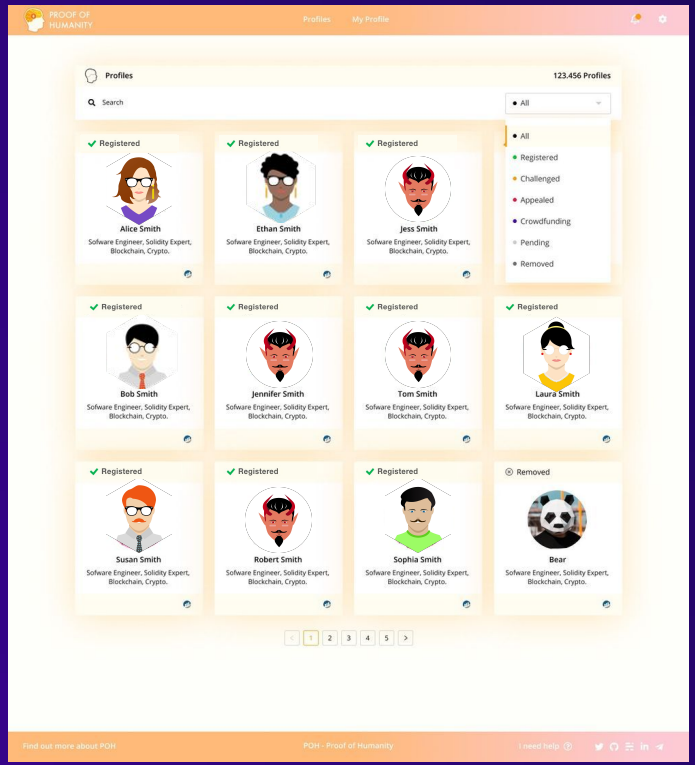
0% overall chance of getting >= 2  **‘s**



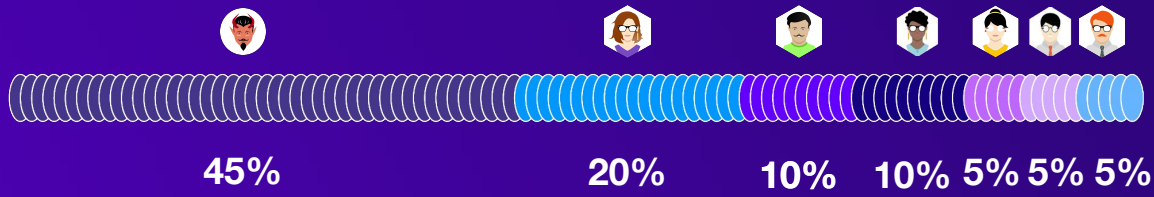
Jury selection with insecure proof-of-personhood



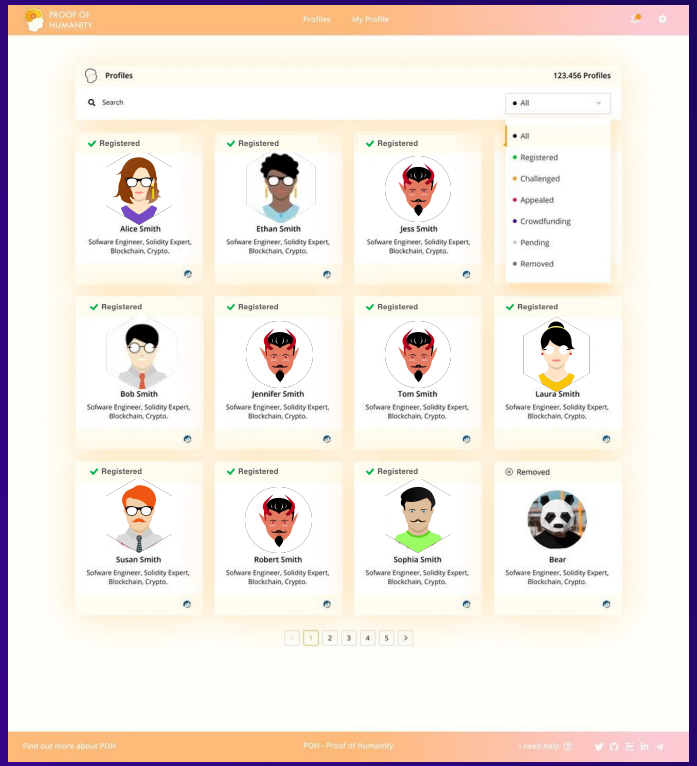
Jury:



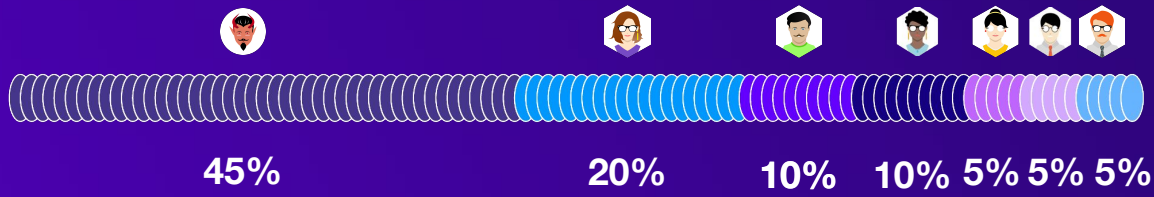
Jury selection with insecure proof-of-personhood



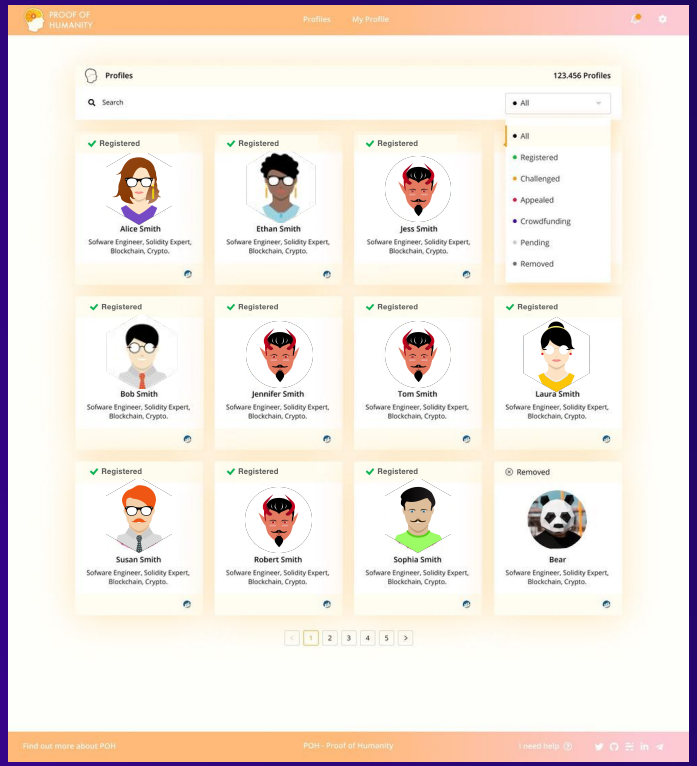
Jury:



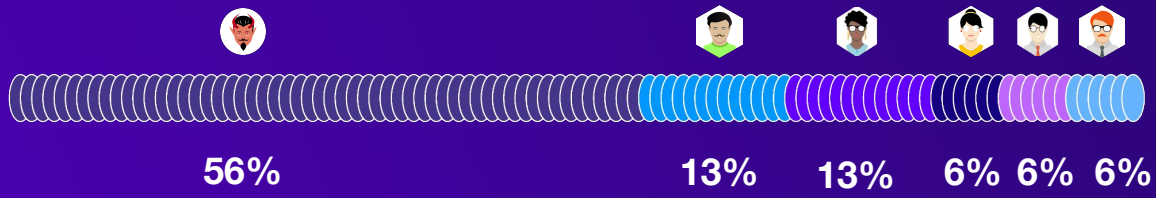
Jury selection with insecure proof-of-personhood



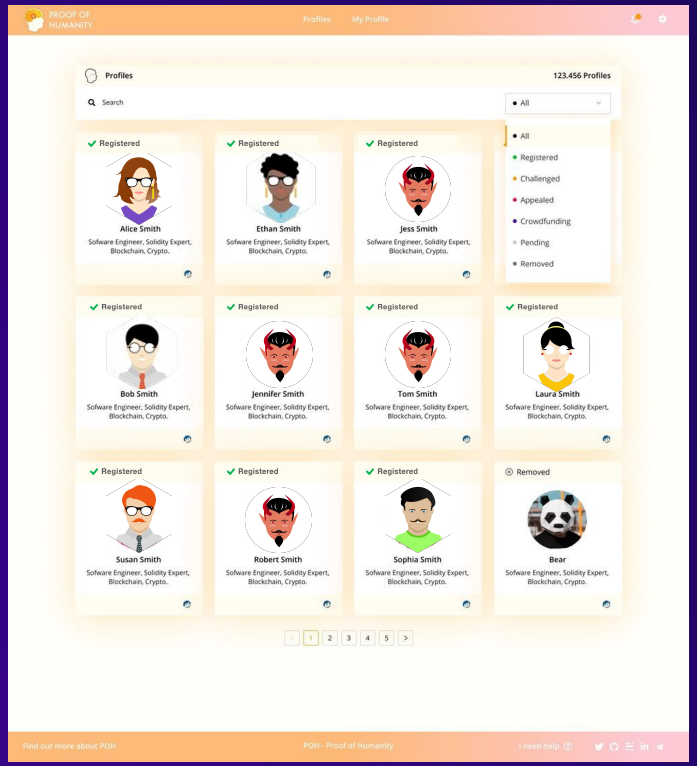
Jury:



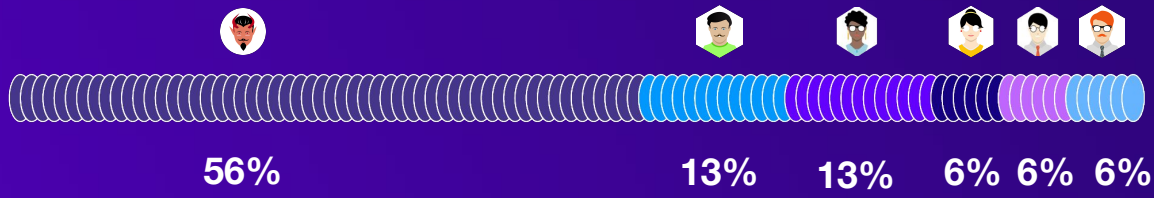
Jury selection with insecure proof-of-personhood



Jury:



Jury selection with insecure proof-of-personhood

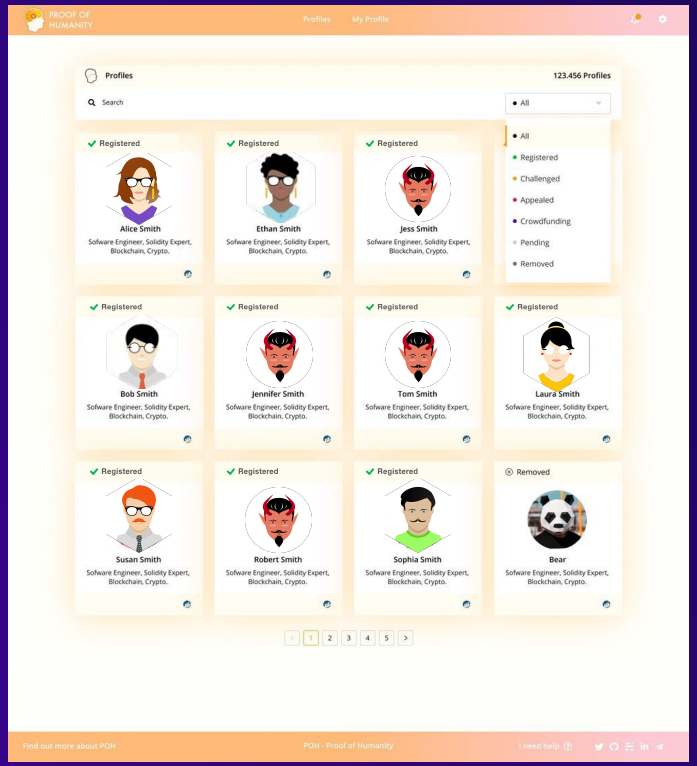


Jury:

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 - in example threshold for  having 50% chance of getting 2 pots is x=44%

Weighting by SBTs

POAP!

Offset Match

(from Decentralized Society: Finding Web3's Soul. Weyl, Ohlhaver, Buterin)

For each pair of participants i,k define:

$$s_{i,k} = \frac{\#\text{SBTs i,k share}}{\#\text{SBTs i has}}$$

Then solve for such that α_i

$$\alpha_i + \sum_{k \neq i} \alpha_k s_{k,i} = 1$$

Offset Weighting

Example:

s , =1/2

α =1/4

α =1/4

α =1

α =1/4

α =0

α =1/4

Offset Weighting - philosophy

$A = (\alpha_1 \ldots \alpha_K)$

$S = (s_{k,j})_{k,j} = 1, \ldots, K$

$1 = (1 \ldots 1)$

GOAL: Solve for A such that $A \cdot S = 1$

Offset Weighting - philosophy

$$A \cdot S = 1$$

$$\Leftrightarrow$$

$$\Leftrightarrow$$

$$\begin{cases} \sum_{j=1}^{K_1} \alpha_j = 1 \\ \sum_{j=K_1+1}^{K_2} \alpha_j = 1 \\ \sum_{j=K_2+1}^{K} \alpha_j = 1 \end{cases}$$

$$(\alpha_1 \ldots \alpha_K) \cdot \begin{pmatrix} 1 & \cdots & 1 & 0 & 0 \\ \vdots & & \vdots & & \\ 1 & \cdots & 1 & & \\ 0 & & & 1\ 1 & 0 \\ & & & 1\ 1 & \\ 0 & & 0 & & 1 \cdots 1 \\ & & & & \vdots \\ & & & & 1 \cdots 1 \end{pmatrix} = (1 \ldots 1)$$

1 to K_1 K_1+1 to K_2 K_2+1 to K

Offset Weighting - calculation issues

- ➔ solution for α_i doesn't always exist, if exists may not be unique
- ➔ May produce negative α_i 's
- ➔ Handles how much weight given to group reasonably well… but not necessarily naturally distributed to individuals. Tendency of offset match to give participants with a superset of SBTs zero weight

Quadratic Funding with Offset Match

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

Then project p receives a grant of :

$$\left(\sum_i \sqrt{\alpha_i c_i^p}\right)^2 - \sum_i c_i^p$$

Juror selection with Offset Weighting











α  =1/4

α  =1

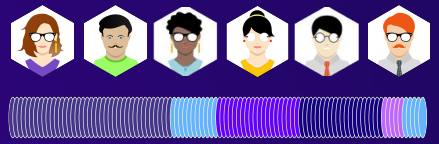
α  =0

α  =1/4

α  =1/4

α  =1/4

Unweighted Stakes



40% 10% 20% 20% 5% 5%

Weighted Stakes

$$\frac{\alpha_i \cdot \text{stake}_i}{\sum_j \alpha_j \cdot \text{stake}_j}$$

36% 36% 0% 18% 5% 5%

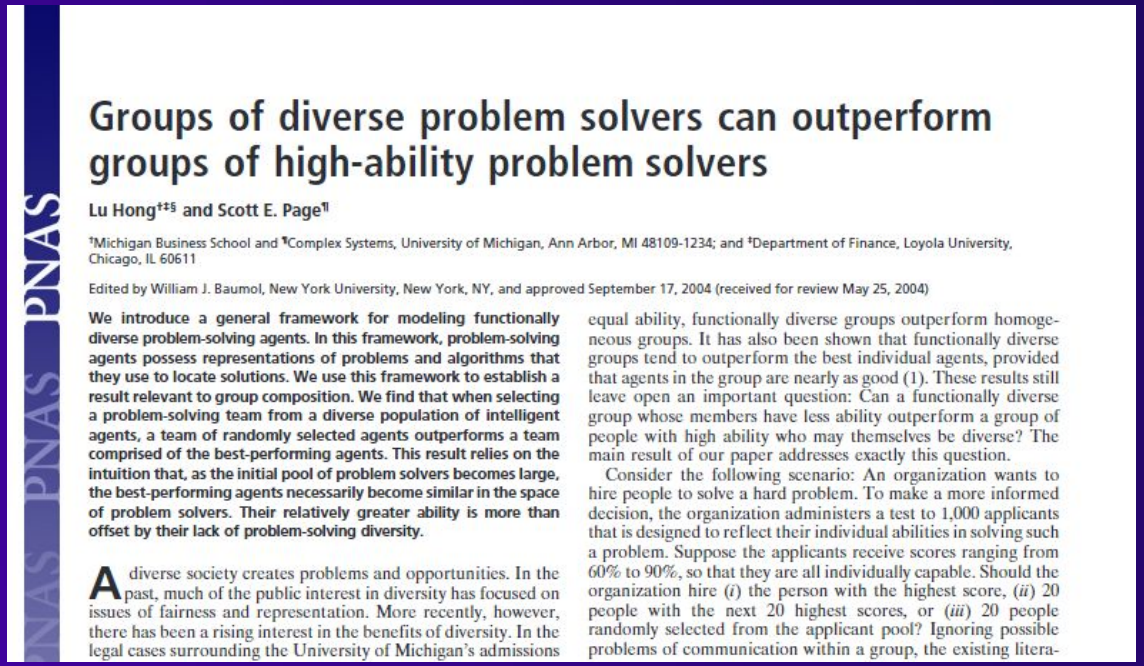
Juror selection with Offset Weighting

Expert in X

Expert in Y

Expert in X

Expert in X



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-

Weighting by SBTs

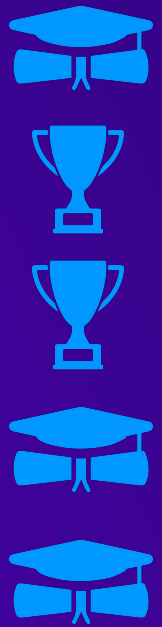
Security Model - Questions to Ask

- ➔ What happens if attacker breaks distribution of an SBT?
- ➔ Are participants incentivized to keep SBTs on same address or split them up?

Weighting by SBTs -  is broken

POAP!

N profiles each with



 only SBT that defines group

Result:  takes almost all weight from  group

Weighting by SBTs -  is broken

 one of K SBTs that characterizes a cluster

 makes N profiles with  can’t copy others

$$(\alpha_1 \ldots \alpha_K) \cdot \begin{pmatrix} \begin{matrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{matrix} & \begin{matrix} 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \end{matrix} & \begin{matrix} 0 \end{matrix} \\ \begin{matrix} \frac{1}{K} & \cdots & \frac{1}{K} \\ \frac{1}{K} & \cdots & \frac{1}{K} \end{matrix} & \begin{matrix} 1 & \frac{K-1}{K} \\ \frac{K-1}{K} & 1 \end{matrix} & \begin{matrix} 0 \end{matrix} \\ \begin{matrix} 0 \end{matrix} & \begin{matrix} 0 \end{matrix} & \begin{matrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{matrix} \end{pmatrix} = (1 \ldots 1)$$

Weighting by SBTs -  is broken

 one of K SBTs that characterizes a cluster

 makes N profiles with  can’t copy others

Result (under offset match):  still takes almost all weight from  group

Other weighting schemes?

Other idea?:

Find $\alpha_1, \ldots, \alpha_K$ with $\alpha_i = \frac{1}{\lambda} \sum_j \frac{1}{s_{i,j}} \alpha_j$

Related to eigenvector centrality/Pagerank algorithm

Conclusions

- ➔ Proof-of-personhood protocols natural substrate for types of social recovery, in some cases distribution of SBTs
- ➔ Ideas developed around weighting QF contributions by SBTs also relevant in juror selection, maybe other applications

Conclusions

- ➔ Moving from:
  purely economic
  - economic + proof-of-personhood
  - economic + proof-of-personhood + SBTs

  can have subtle effects on security model
- ➔ Questions relevant to security models of systems using SBTs - what happens if a single SBT broken? Do users have an incentive to split SBTs over different wallets?

Economic Incentives and Souls in Schelling-point Based Oracles

WILLIAM GEORGE
Research Lead, Kleros

kleros.io

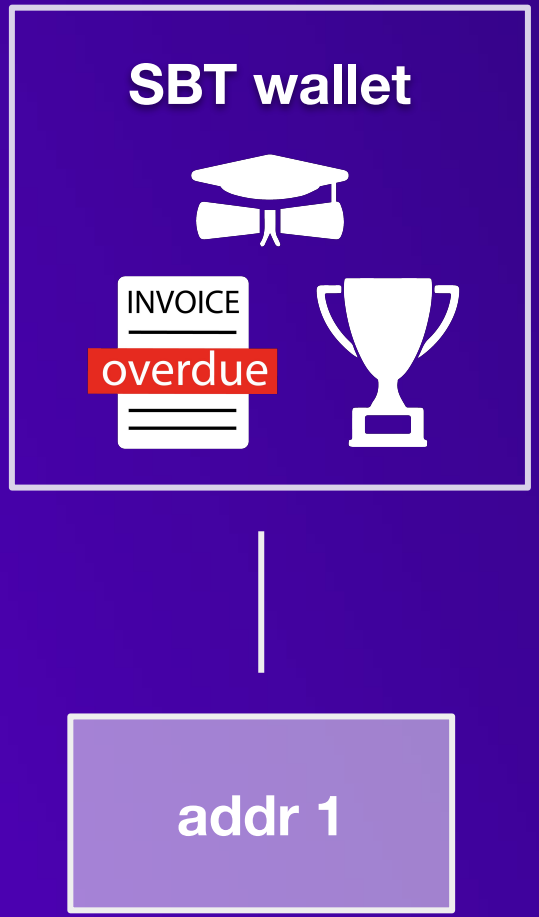
KLEROS

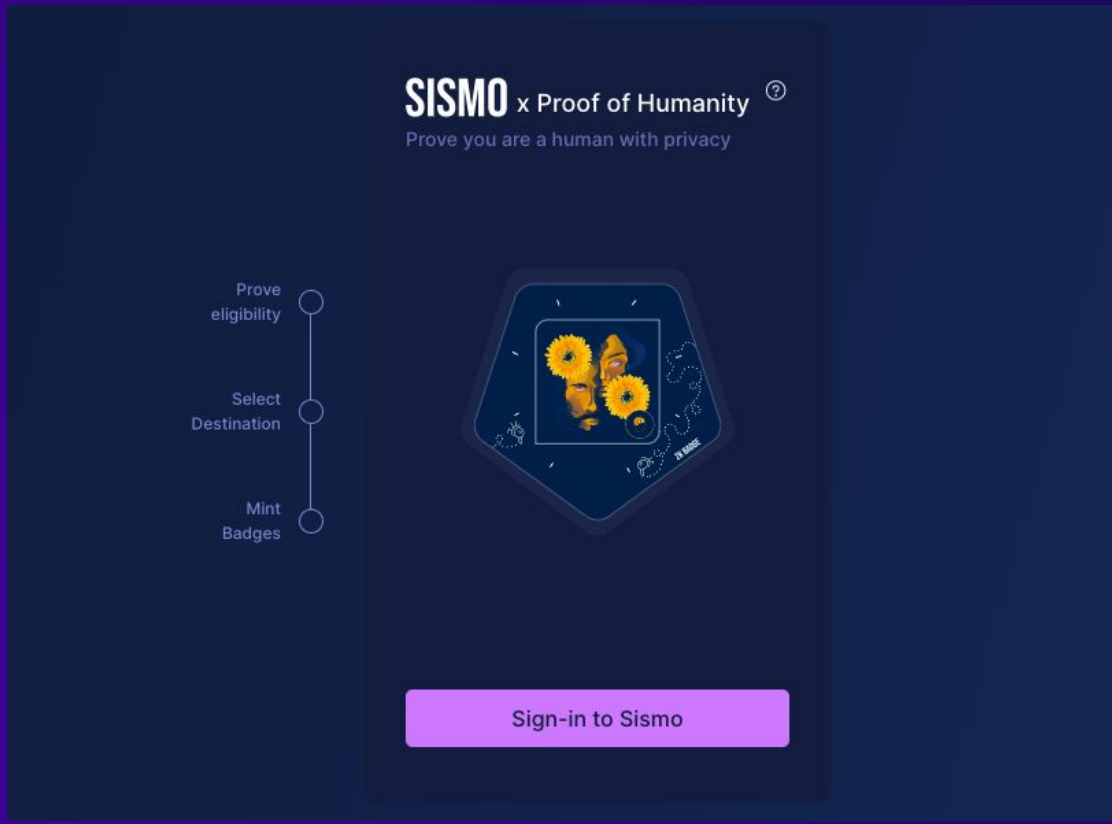
HOSTED BY
THOMSON REUTERS
INCUBATOR

European Commission

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

bpifrance


SBT wallet
INVOICE
overdue
addr 1


SISMO x Proof of Humanity
Prove you are a human with privacy
Prove eligibility
Select Destination
Mint Badges
Sign-in to Sismo