Personhood Credentials

From Theory To Practice

Steve McCown

Anonyome Labs, Chief Architect
DIF, Steering Committee
9 May 2025

I'll Believe It … When I See It?

Is this real or fake?

How do you know?

I'll Believe It … When I See It?

How about this one?

Now It Gets Serious…

CNN: Finance worker pays out $25 million after video call with deepfake ‘chief financial officer’

“…worker had grown suspicious after he received a message that was purportedly from the company’s UK-based chief financial officer.”

“…the worker put aside his early doubts after the video call because other people in attendance had looked and sounded just like colleagues he recognized.”

...and Potentially Fatal

26.11.2024

Breakthrough in the treatment of type 2 diabetes in Australia

- Real Doctor
- Deepfake Video
- Promoted on Facebook
- Stop taking Metformin
- Try “All-Natural Drug”

The AI Problem … is Growing

AI makes fraud easy, fast, and cheap.

Solution: Personhood Credentials

100K Foot View:

1. Give everyone a digital credential
2. Digitally sign everything
3. Make it super easy to verify

Plus:

Must preserve personal privacy … and not track users

What's a PHC?

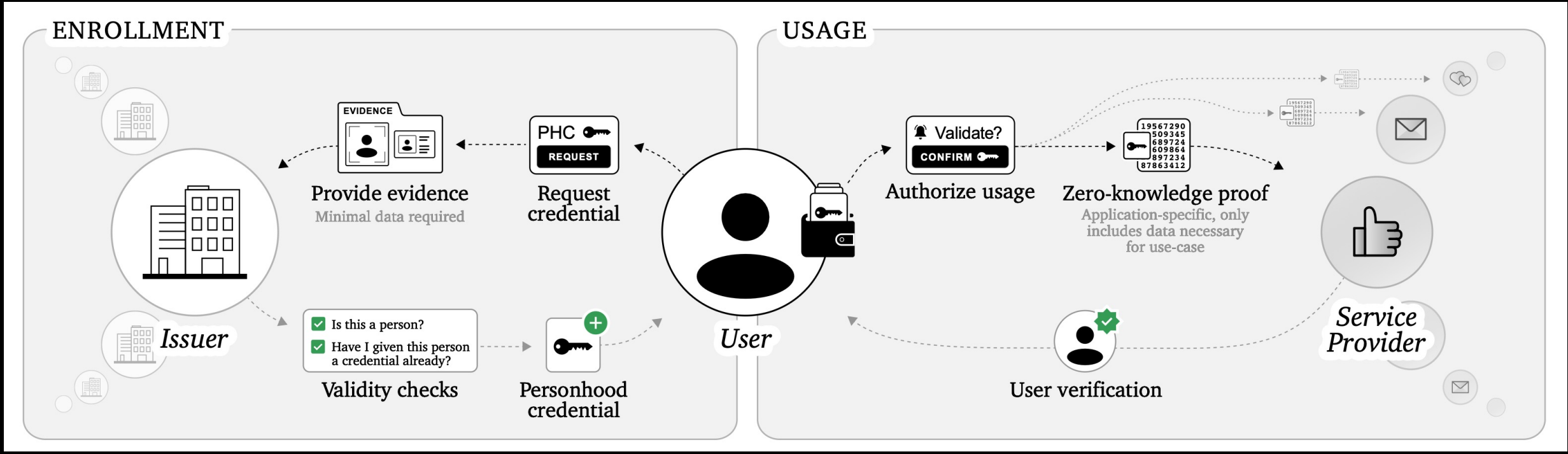
Digital Credential that

- Empowers holders to prove they are a real person … and nothing more

Foundational Requirements

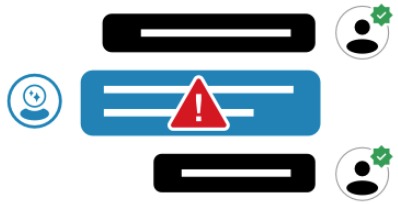
1. You only get 1
2. Unlinkable Pseudonymity
   - Anonymous interaction through a service-specific pseudonym (e.g., handle)
   - Untraceable by the Issuer
   - Unlinkable by providers
3. Not forgeable by AI

Enrolling and Using PHCs



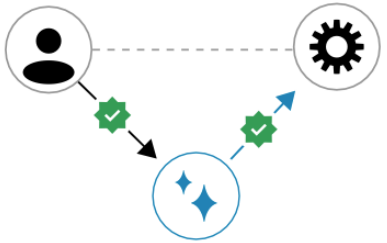
Issuers: governments, schools, hospitals, etc.

Key Benefits



Reduce impact of sockpuppeting

Mitigate bot attacks



Verify delegation to AI agents

Potential Challenges

- **Equitable access:** can PHCs affect access to digital services?
- **Free expression:** will people feel safe using PHCs?
- **Checks on power:** how will PHCs affect tech providers?
- **Robustness to attack and error:** how might PHCs be vulnerable to errors or subversive compromise?

Technical Cautions

- Crypto
  - Quantum Safe Crypto
  - Signatures must support Zero Knowledge Proofs
- Prioritize Privacy and Security *above* Interoperability
- NO “Phone Home” requirement to verify credentials

Future Considerations

Agentic AI

- Entrust your credential to an AI operator

Wait … aren't we trying to detect / block AIs?

Well, yes … but people will still want to *use* them!

Further Reading

Paper

- Personhood credentials: Artificial intelligence and the value of privacy-preserving tools to distinguish who is real online
- https://arxiv.org/abs/2408.07892

Thank You!

Steve McCown

~ Email: smccown@anonyome.com

~ Discord: @mccown

~ LinkedIn: https://www.linkedin.com/in/mccown

ANONYOME LABS

https://anonyome.com

DIF

https://identity.foundation