Verifying Humanness: Personhood Credentials for the Digital Identity Crisis

Ayae Ide

Tanusree Sharma

Presenter: Tanusree Sharma
Assistant Professor
Penn State

PennState

Agenda

1. Problem
2. Policy
3. Consumer Insights
4. Design Recommendation

Identity Fraud, Bots, Fake Accounts

Survey Bots Exposed: How They Manipulate Data and Skew Results

Sanja Trajcheva | Marketing | October 17, 2023

PERSONAL FINANCE

Social Security Administration updates information on new anti-fraud measures for benefit claims

PUBLISHED WED, APR 9 2025·3:16 PM EDT | UPDATED THU, APR 10 2025·8:32 AM EDT

Lorie Konish

SHARE

KEY POINTS

- The Social Security Administration provided more information on new identity proofing policies set to go into effect on Monday.
- The agency will now continue to allow more types of beneficiary applicants to apply by telephone. However, those who are flagged will be required to visit an agency office in person.

Borys Musielak @ Warsaw
@michuk · **Follow**

You can now generate fake passports with GPT-4o.

It took me 5 minutes to create a replica of my own passport that most automated KYC systems would likely accept without blinking.

The implications are obvious --any verification flow relying on images as "proof" is now officially Show more

10:33 AM · Apr 1, 2025

https://securityaffairs.com/176224/security/chatgpt-4o-to-create-a-replica-of-his-passport-in-just-five-minutes.html
https://www.bbc.com/news/articles/c4ng24pxkelo

Traditional & Emerging Mechanisms to verify Human or Human Interaction

Heuristics, Behavioral
Guixin et.al., Meriem et.al., Andrew et.al.

ID.me

Biometrics, Physical IDs, Video with Trusted referee

Decentralized Identifier (DID) Wiki

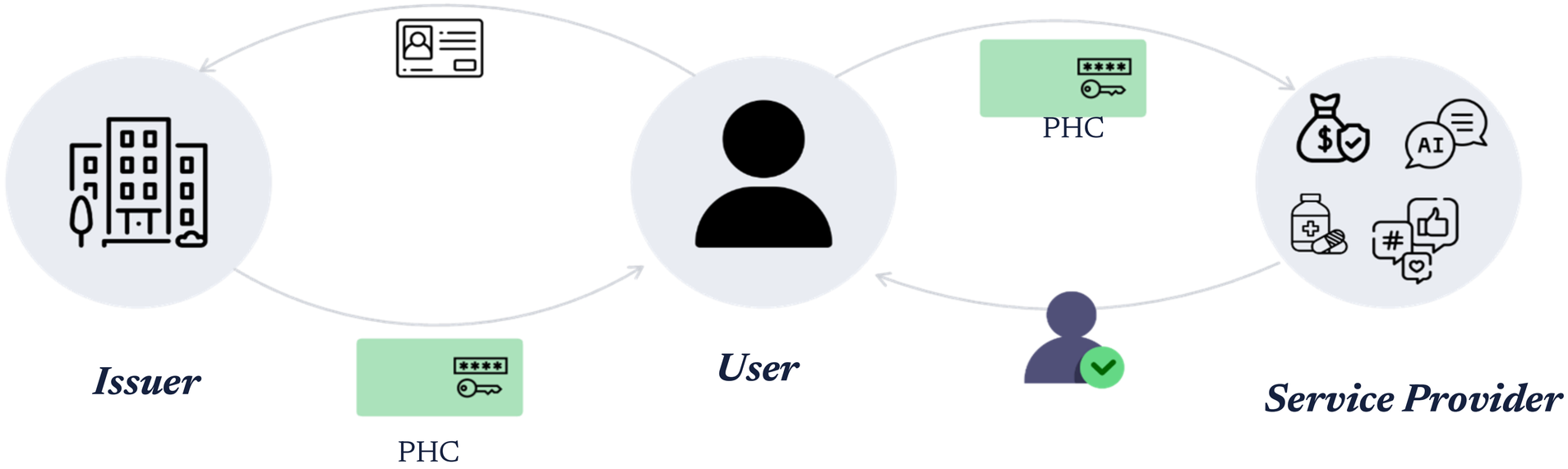
W3C Decentralized Identifier Working Group

https://www.dock.io/post/why-amazon-is-doubling-down-on-digital-credentials-mdls-and-eudi-video-takeaways

Personhood Credentials (PHCs): empower users to demonstrate to services that they are real people, without disclosing personal information

1. Request credential with verification evidence

3. Authorize usage with PHC via zero-knowledge proof



2. Issue Personhood Credential

4. Verify User

Emerging Mechanisms to verify Human

https://www.wired.com/story/sam-altman-orb-store-san-francisco/

What **factor influence users' preferences** on how they would like to verify themselves?

Personhood Credentials → **Needs** **Contexts** **Preferences** …

1. Competitor & Cognitive Walkthrough (n: 12)

Identify common patterns and gaps of existing verification systems

2. Policy Analysis (n: 18)

Regulatory constraints of identity systems

3. **User Study + Design Sessions (n: 27)**

Factors & Design

Policy Overview of Digital Identity System Requirements

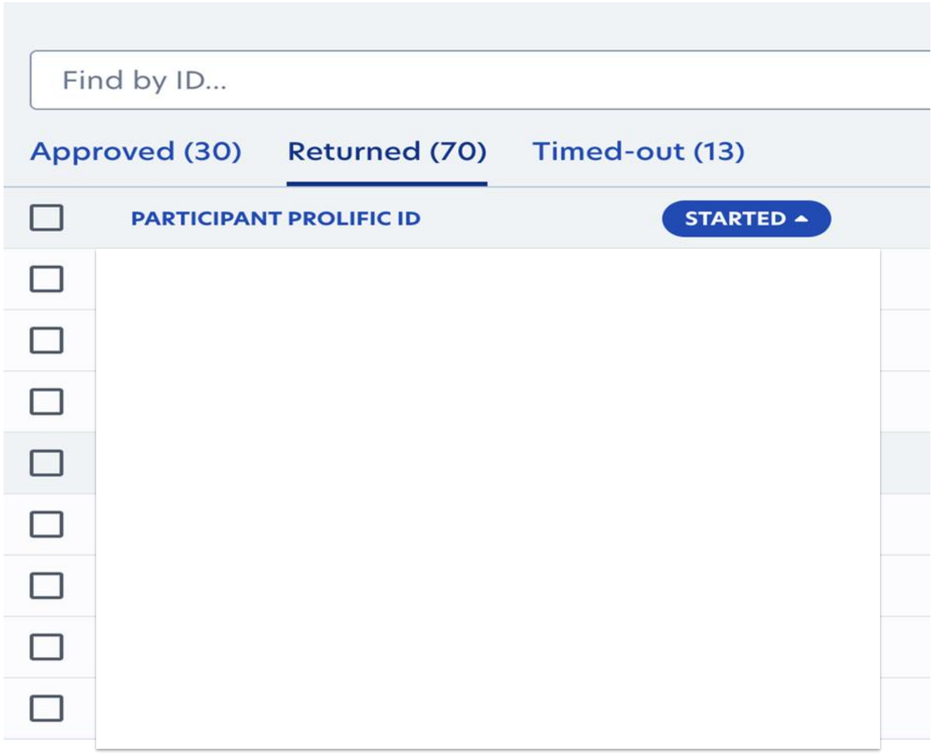
	eIDAS	California Digital ID Framework	NIST Digital Identity Guidelines	W3C Recommendation
Data Requirements	Not clearly defined yet. (EU Digital Identity Wallets is still under testing phase.)	**One or more of the following**: • Something you know (e.g., password) • Something you have (e.g., one-time code) • Something you are (e.g., face)	Data requirements of identity proofing (e.g., ID documents, biometrics) **based on Assurance Level** (e.g., IAL2, IAL3)	Verifiable credential can contain any data attested to by an issuer, allowing for **Selective Disclosure.**
Data Storage Requirements	• **Decentralized**: Data is stored on user's device wallet • Member states should integrate **privacy-preserving technologies**. (e.g., ZKPs)	**Centralized**: California Identity Gateway minimize ID and eligibility data and share to third parties.	Data protection **requirements based on Assurance Level** (e.g., Approved cryptographic techniques are required at IAL2.)	Prove that the information is not tampered with a **cryptographic proof** (e.g., digital signatures and zero-knowledge proofs.
Issuer Requirements	**Trusted organisations** that have been authorized by a national competent authority.	Identity Gateway to federate identity providers and agencies.	Credential service providers and identity providers must meet NIST requirements.	Any entity can be an issuer of Verifiable Credentials if they can **cryptographically sign** the data.

Project-page: https://shorturl.at/XNMox

Impression: Fairness, Unknown Privacy Guarantee

[Fairness & Representation] often time I got "**time out or returned**" in this platform too often lately

[Unknown Privacy] **I would still stay with email or traditional verification as I know what they are keeping**, I don't exactly know how much privacy I have in human tests.



Prove that you are a real and unique person, without revealing who you are.

World ID is private by design, offering you unmatched control over your information. Once you verify your World ID, you can use it to prove that you are a real and unique human without revealing who you are.

This is made possible by a cryptographic technology called Zero-Knowledge Proofs (ZKPs), which can prove both that your World ID is real and that you have never done the action you are trying to do before – all without ever revealing who you are.

With World ID, we rethink internet identity. It does not matter who you are, just that you are a unique human (proof of human). This will become increasingly relevant with the rise of artificial intelligence and more sophisticated bots that behave like humans.

USER INSIGHTS

Factors: PHC Issuer

Most Trusted Issuer: Government

Context-Dependent Preferences

Context	Preferred Issuers	Reasoning from users
If verification needed for government services..	Government, Local Authorities	{*Netherland, Estonia's ID issuance are based on municipalty, so I would imagine* ***same issuer from local government:*** "Familiarity"} {I feel like **central authority can respond to any threat aftermath** if something goes wrong "Effective Reactive"}
If verification needed for onboarding to socials....	Private companies with oversights	{I don't want gov-issued PHC for social media verification.. I don't want another layer of surveillance **just to log into Instagram**: "Surveillance"}
If verification needed for health claim, financial claim.....	Npo with government Oversights Industry (healthcare, finance) tied to legal themselves	{Not as intense as the government: "Balanced trust and flexibility" *I trust my bank because* ***they are regulated by HIPPA****. I can trust them as issuer too. "Regulatory ties"*}

USER INSIGHTS

Factors: Data Types for PHC Issuance

Minimize Data Sharing

Government id (**multiple info**) vs Biometric (**single info**)

Sensitivity, Security, & Efficiency Across Biometrics

Fingerprint is fine, but **face** is too much.
Iris verification is probably the more secure of all, fingerprints and facial recognition can be regenerated
Face is more technology improved.

Source: https://goldenson.com/2016/11/28/startup-tech-to-get-government-id/
Source: https://www.vecteezy.com/vector-art/20897439-biometric-security-of-fingerprints-facial-iris-glyph-style-icon-set-editable-vector-eps10

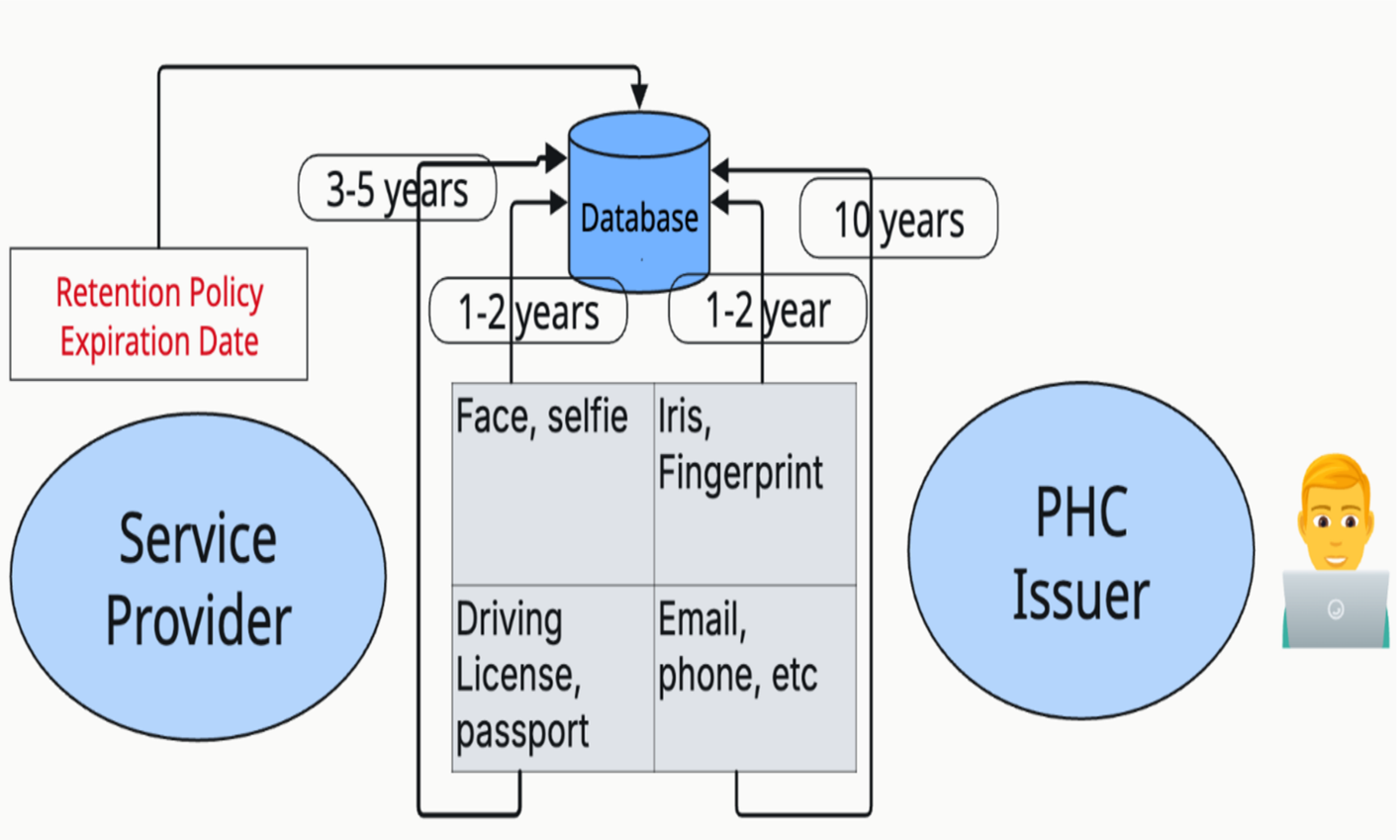
USER INSIGHTS

Design for PHC: Time Bound, Periodic Biometric

Time-Bounded Credentials

Time-bounded credential with retention and expiration date based on different data types

- Proof without storage, pseudonymization
- Design for Delegated Onboarding without compromise
- Dynamic authentication



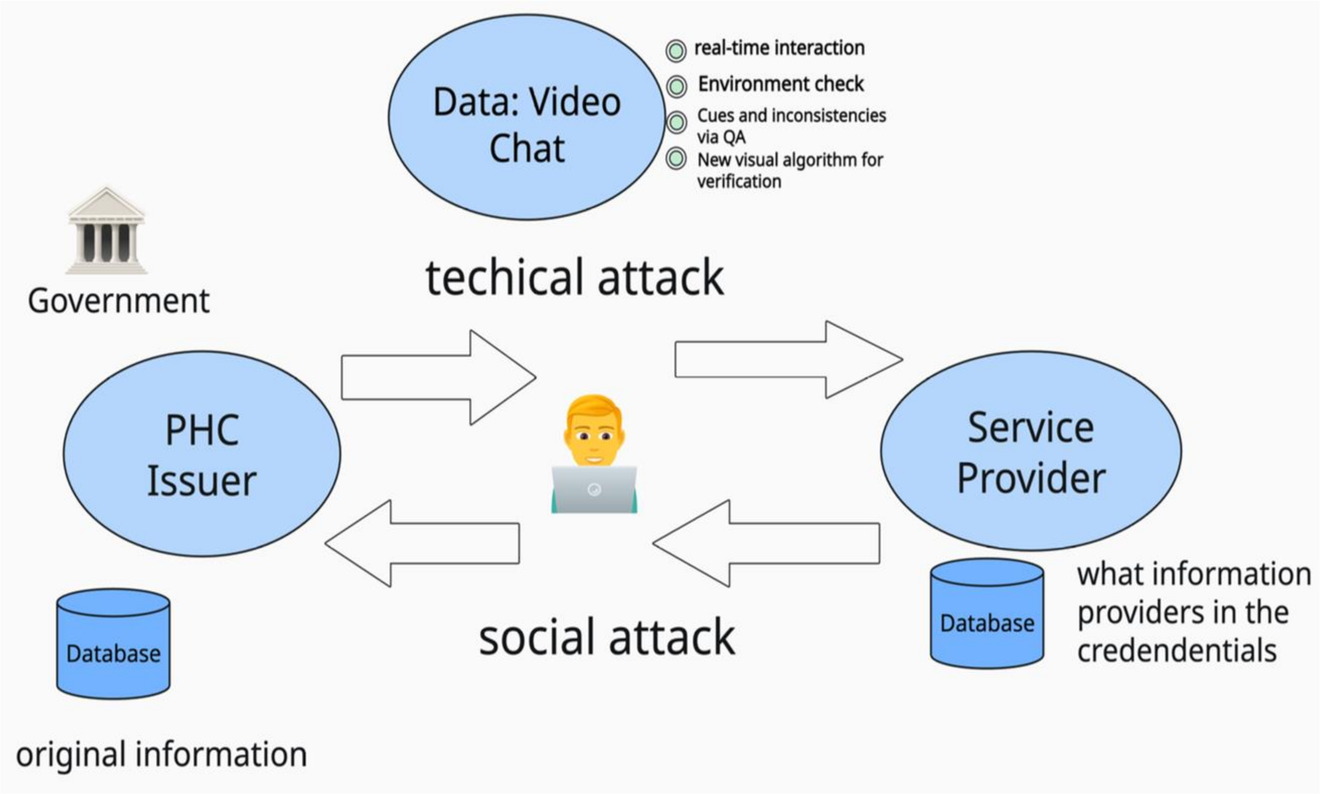
USER INSIGHTS

Design for PHC: Multi-Factor Humanness

Multi-Factor Human Check

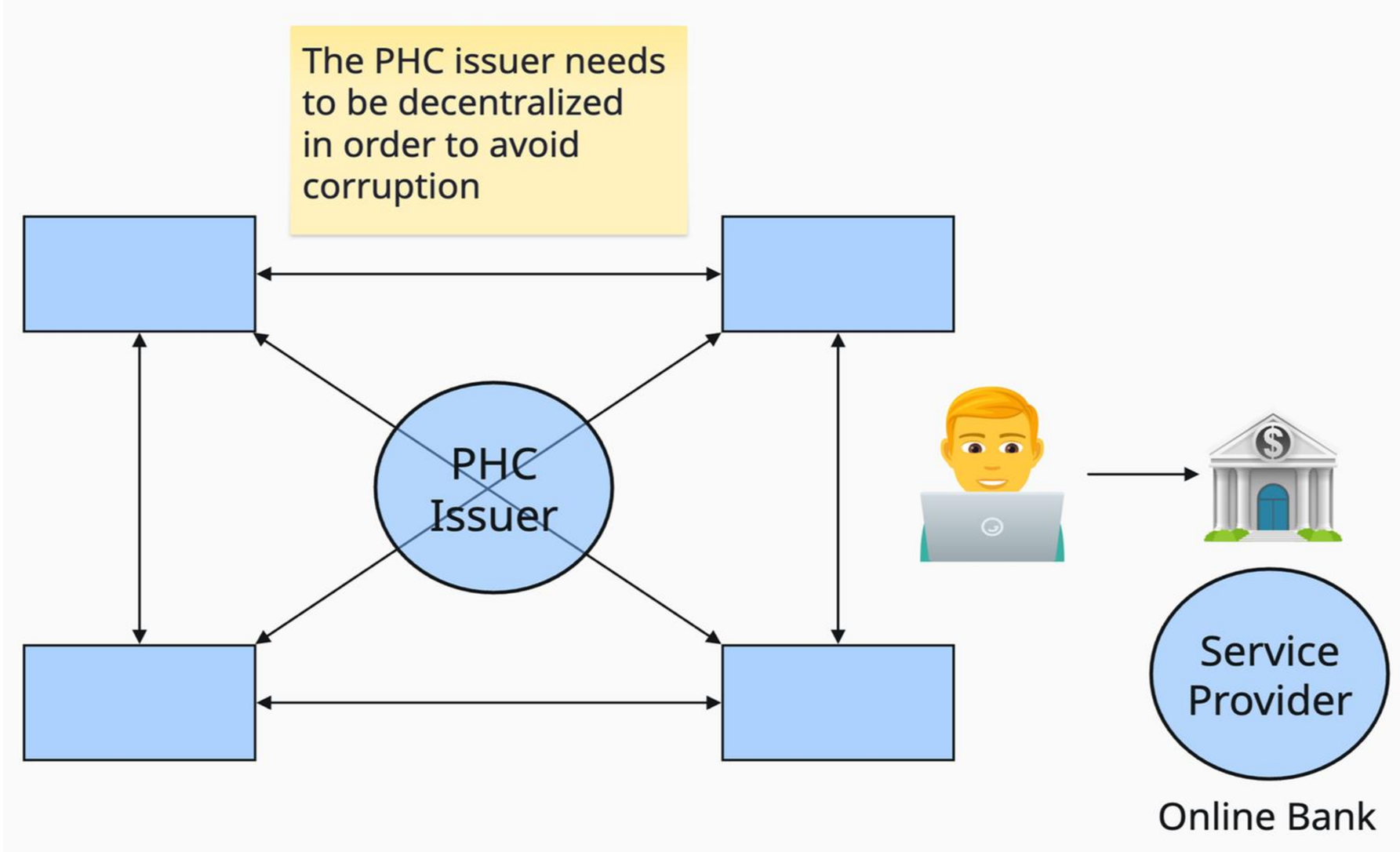
Visually interactive human check with video chat for humanness cues, environment check

- Low latency video at Scale (reasonably with secure end-to-end encryption)
- Highly Trained Vetted Issuers
- Built-in Environment Analysis.



USER INSIGHTS

Design for PHC: Distribute Power Across Issuers



Takeaways

- How do we build trust & scale personhood credentials?
- Standardization of personhood credential
  - Issuer
  - Data
  - Issuance system
  - Required / suggested technology stacks for engineers

Project Page

Contact: tanusree.sharma@psu.edu