Age and Identity Verification: A Review

An ongoing open-source literature review posted and curated by Jonathan Haidt (NYU-Stern) and Zach Rausch.

This Google doc is a working document that contains the citations and links to articles, essays, and organizations whose work is relevant to creating fast and reliable ways to authenticate the age and/or identity of people who want to open an account (e.g., on a social media platform) or access an age-restricted service (such as pornography or gambling).

The document is split into four sections. After a brief introduction, we examine identity verification methods, then age verification methods.

If you are a researcher and would like to notify us about other studies, or add comments or counterpoints to this document, please request access to the Google Doc, or contact Haidt directly, and he will set your permissions to add comments to the Google doc. This document is evolving based on feedback.

Thanks to Cedric Warny, Camille Carlton (Center for Humane Technology), xxx for suggestions to improve this document.

Notes:

- Also see our companion reviews:
  - Is there an increase in adolescent mood disorders, self-harm, and suicide since 2010 in the USA and UK? A review [with Jean Twenge]
  - Social media and mental health: A collaborative review [with Jean Twenge]
  - Social media and political dysfunction: A collaborative review [with Chris Bail]
- See also additional Google docs laying out evidence for trends in mental health and social media use in Australia, Canada, New Zealand, and other countries.
- Last updated: September 18, 2023

* * * * * * * * * * * * * * * * *

Clickable Table of Contents

1. Theoretical and Legal Issues

1.1 On the need for and legitimacy of federal regulation

BRIEF INTRO: xyz…

1.1.1 Jones & Samples (forthcoming 2022). On the Systemic Importance of Digital Platforms. *University of Pennsylvania Journal of Business Law.* (h/t Tim Samples)

CONDENSED ABSTRACT FROM TIM SAMPLES: Proposes a theoretical basis for imposing a prudential regulatory regime for digital platforms based on their systemic importance, drawing parallels with the framework for systemically important financial institutions (SIFIs) in the Dodd-Frank Act.

1.1.2 Werbach & Zaring (forthcoming 2022). Systemically Important Technology. *Texas Law Review*. (h/t Tim Samples)

CONDENSED ABSTRACT FROM TIM SAMPLES: This article addresses the risks of failure within the connective tissue of systemically important network institutions.

1.1.3 Griffin (forthcoming 2021). Systemically Important Platforms, *Cornell Law Review*. (h/t Tim Samples)

CONDENSED ABSTRACT FROM TIM SAMPLES: This article proposes a special designation for systemically important platforms centered on their use of manipulative technologies.

1.1.4 Öhman & Aggarwal (2020). What if Facebook Goes Down? Ethical and Legal Considerations for the Demise of Big Tech. *Internet Policy Review*.

CONDENSED ABSTRACT FROM TIM SAMPLES: This article explores the failure risks of Facebook, coins the term systemically important technological institutions (SITIs), and proposes more research in that area.

[Other studies? What have we missed?]

1.2 Privacy Concerns

[Others? What have we missed?]

* * * * * * * * * * * * * * * * * *

2. User Authentication

One of the main reasons that social media platforms are toxic to democracy is that they are a gift to trolls, Russian intelligence agents, political operatives, swindlers, and anyone else acting in bad faith who can create one or thousands of accounts. Many reform proposals (including those from Elon Musk, Jonathan Haidt, Jamie Dimon, …) talk about the benefits of requiring some form of user authentication. But what does that mean? First, it is crucial to note that authentication does NOT mean that people must post using their real names. Rather, under most authentication schemes, anyone can still open an account, instantly, on platforms such as Facebook or Twitter, with a pseudonym and no authentication, if they simply want to view the posts of others. But then, as a second step, for those who want to post their own content and gain algorithmic amplification to a potentially vast audience, users would be required to take a subsequent step of authentication, likely carried out by a 3rd party company or non-profit. There are (at least) three levels of authentication.

Level 0 = No authentication.

This is what we have now. Any person or automated system can create unlimited fake accounts every day.

Level 1 = authenticate humans

Users must pass a captcha, to show that they are a human and not a bot. But each human could still create and run hundreds of troll accounts, or create them and turn them over to AI to run.

Level 2 = authenticate unique identity once and untraceably.

This would be carried out by a non-profit or for-profit company, using a variety of methods. A user at Facebook (for example) who wants to be able to post would get sent over to this third party. Any methods that require showing a government ID, or giving biometric information, would then wipe out the information after authentication when sending back the approval to the platform requesting authentication. These schemes allow each person to create only one account. Examples of companies or non-profits who are developing such schemes:

Level 3 = authenticate identity to a 3rd party, who keeps the information.

3.1 A company like Clear is well situated to do this, as it already does for air travel, sporting events, and many other situations where there is a need for security balanced with privacy.

3.2 India's Aadhar platform authenticates people in real-time. Aadhar stores encrypted biometric data. Aadhar is maintained by "The Unique Identification Authority of India (UIDAI).

3.3 Human-id.org (hashing solution)

3.4 World Coin: see Vitalik's discussion (Hashing solution)

3.5 Proofofexistence.xyz (Hashing solution)

3.6 Soulbound tokens: blogpost, paper, Kate Sills' critique (Hashing solution)

3.7 BrightID (Vouching Solution)

2.6 Proof of humanity (Vouching solution)

Additional Notes

Question 1: What about protecting dissidents in repressive countries?

Answer: Why does the whole world need to be on a single platform? That was a dream ten years ago, but now it appears that we might need one kind of platform optimized for the "public square" of advanced or stable democracies, with incentives for constructive dialogue, and a very different set of platforms designed for life in the more dangerous "public square" of authoritarian countries, where the design imperative is for untraceability and protection of dissidents. It would be trivially easy to connect the two platforms: journalists or human rights organizations on the democratic platforms can simply re-post content from dissidents and whistleblowers on the high security platforms, without even knowing their real identities.

Question 2: What about whistle blowers or political groups who want a second account? Is everyone limited to one authenticated account?

Answer: There would be provisions for accounts beyond the regular single-person accounts. Companies and non-profit organizations would certainly have accounts, and there would be provisions for authenticating them. Whistle blowers would still have hundreds of ways to get news out to the world, anonymously, via blogs, journalists, anonymous hotlines, and non-profit accounts that could be set up for the purpose. It's not clear why critics and whistleblowers must each have their own individual un-authenticated Twitter or Instagram account to be effective.

To learn more about user authentication

- See this essay by Scott Galloway, on the necessity of identification in the online world
- Listen to this episode of Brave New World, a conversation between Vasant Dhar and Jonathan Haidt. (Discussion of KYC is towards the end of the episode).
- Tom Newton Dunn: We must bite the bullet on online anonymity to defeat the trolls (*Evening Standard)*.

[Other studies? What have we missed?]

* * * * * * * * * * * * * * * * * *

3. Age Gating and Age Verification

First, read the history of How 13 Became the Internet's Age of Adulthood, back in 1998. It was supposed to be 16, but lobbyists for e-commerce companies got it lowered. There was no consideration of mental health; this was about when children can sign contracts with companies to give away their data and their rights, without any parental permission. 25 years later, the internet is very different and studies show that young teens (11-15) are the most badly harmed by spending time on social media. The age should be raised, but how to enforce it, rather than relying on the honor system as we do today? Jon Haidt suggests that companies that need to enforce a minimum age should be required to offer a menu of methods by which customers could prove that they were old enough, rapidly and reliably. One option can be posing for a selfie with one's driver's license or other government-issued ID, as some companies do now, but there are so many other ways, for people who do not want to share their ID, or even their real name, with the platform. For example:

3.1 There are already many companies devoted to checking the age of potential customers, rapidly and conveniently. There are so many of them now that they have their own trade association: The Age Verification Providers Association. Examples include AgeChecker.net, or Yoti.

3.2 Clear (which you know from airports) already handles age verification rapidly and conveniently, e.g, for customers who want to buy beer at sporting events.

3.3 See multiple proposals here: Chris Griswold (2022) Protecting Children from Social Media. National Affairs. E.g.: "One possibility would be for the SSA [Social Security Administration] to offer a service through which an American could type his Social Security number into a secure federal website and receive a temporary, anonymized code via email or text, like the dual-authentication methods already in widespread use. Providing this code to an online platform could allow it to confirm instantly with the SSA whether the user exceeds a certain age without further personal data reaching the platform or the government."

3.4 See Yuval Levin's NYT essay: How Changing One Law Could Protect Kids From Social Media.

3.5 Facebook developing AI, new ways to detect users under age 13.

3.6 See the UK Age appropriate design code. See also Age Verification: State of Play and Key Developments in the EU and UK. The UK issued s series of age verification recommendations for gaming.

3.7 Meta is testing a new age verification system, offering users three ways to prove they are the age they say they are. BUT: it seems that they only do this if a user tries to change her age to make herself older. If users lie about their age when they create the account, they are OK.

[Other studies? What have we missed?]

* * * * * * * * * * * * * * * * *

4. World Wide Web Consortium (W3C)'s Vision

Terms: DID: decentralized identifiers; VC: verifiable credentials; ZKP: zero-knowledge proofs; Digital wallets / agents.

- VC standard: https://www.w3.org/TR/vc-data-model/
- DID standard: https://www.w3.org/TR/did-core/
- The VC standard supports zero-knowledge proofs: https://www.w3.org/TR/vc-data-model/#zero-knowledge-proofs

The DID/VC standards are ecosystem-agnostic, and backed by the top researchers in the field of digital identity. That said, few real-world products currently exist that implements those standards. There are various reasons for this, including the fact that this is fairly new, some of the standards are still WIP, and there's been some resistance by big tech names like Google. That said, there are a couple of high-profile projects that make use of this tech:

- The company behind USDC (the biggest stablecoin) is building a digital identity infrastructure leverage DID/VC: https://www.centre.io/verite
- Block, Jack Dorsey's crypto company, is working on a set of technologies which they facetiously call "web 5". VC/DID are a central part of this tech stack. Their flagship product is called TBDex, which is a protocol to exchange fiat and crypto in a decentralized manner (as opposed to going through a central node like Coinbase). It's still in development afaik, but a very exciting project.
- Gitcoin Passport: uses DID on Ceramic (a special-purpose blockchain) and VCs (only Gitcoin-issued VCs for now). I'm concerned though that the VCs may be stored on Ceramic itself, which goes against the design I advocate for (need to double-check that). This is probably the closest thing to the soulbound token vision, but leveraging the DID/VC standards. It's like a hybrid between W3C's vision and the soulbound token vision.

You can call this vision the “web5” vision. Also see Cedric Warny’s post on online identification.

Privacy Spectrum

To represent the spectrum of surveillance as described, we can create a visual diagram that places "Ultimate Privacy" at one end and "Dystopian Surveillance" at the other. Along this spectrum, we'll denote the increasing levels of information required by the service provider and the corresponding increase in surveillance capabilities. Here's a basic outline:

Ultimate Privacy:
- Only proof of meeting a basic criterion is needed (e.g., age threshold).
- Uses Zero-Knowledge Proofs (ZKP).
- Assumes private keys remain private.

Moderate Surveillance:
- More data points are required for verification.
- Examples: Actual age, location, picture.
- Digital signature from parents or guardians.

High Surveillance:
- Richer sign-up data.
- Random checks at place of residence.
- AI technology to match user behavior with sign-up claims.

Intense Surveillance:
- AI chatbot calls to parents or guardians.
- Constant monitoring of behavior patterns.

Dystopian Surveillance:
- Maximum data collection.
- Full-scale monitoring and continuous verification.

This spectrum will also highlight the key points mentioned:
- Importance of private keys and the assumption of their privacy.
- Evolution of social/cultural norms around private keys.
- Technological solutions to potential issues.