Exploring Privacy in ID-Based Age Verification Architectures

Sarah Scheffler and Shuang Liu

Carnegie Mellon University

August 2025

Abstract

This submission argues for the following points, gleaned from our team's ongoing research in applied cryptography and policy analysis of age verification based on identity documents (IDs):

1. Standards should prioritize age verification methods that transmit less information, share sensitive information with fewer entities, and allow users more control over which entities to share sensitive information with;
2. The most privacy-preserving methods for ID-based age checking that maintain verifiability are *trusted notaries* (often called *selective disclosure* or *trusted attestation* in the context of age or credential verification) or *cryptographic zero-knowledge proofs* – we discuss these in the body;
3. Sensitive private information is not limited to the content of the ID;
4. When presenting credentials in a privacy-preserving way, holders should know what information they are sending;
5. Standards should take a broad definition of "identifying information"; and
6. Standards should encourage purpose restriction and deletion of information collected during age verification as a best practice.

Introduction

Many jurisdictions, including twenty-five U.S. States, have recently enacted new age assurance policies for online content, requiring certain content providers and websites to take steps to ensure that parties accessing sensitive content are above some minimum age. Methods for estimating or verifying age include biometrics, behavioral signals, action/account history, or identity documents (IDs).

The ID-based approach warrants particular scrutiny. Beyond establishing identity, IDs can also be used to confirm non-identity attributes such as age, residency, or status as a "real" or legal person (i.e. "not a robot"). However, naive ID-based verification carries significant privacy costs and increases the risk of identity theft, since this method encourages widespread transmission and collection of prototypical Personally Identifiable Information (PII). At the same time, the creation of new standards for private ID-based age verification presents an opportunity to build privacy-respecting infrastructure for these checks.

This submission draws on insights gleaned from our year-long investigation into cryptographic privacy-preserving ID-based credential verification, building on preliminary work presented at the 9th Workshop on Technology and Consumer Protection [7] and ongoing research.

Modeling Assumptions

We model a system in which a credential/ID issued by a credential/ID *Issuer* is held by a held by a *User*,[1] who engages in a protocol with a *Gatekeeper* who makes a decision on whether or not to admit the user to some age-restricted website or content. In other words the Gatekeeper wishes to verify whether the age on the User's credential/ID is above a threshold. In some of the models, the User also has a program on their

[1] Although it is sometimes helpful to distinguish between the *User* of an age verification system and the *Holder* of an ID credential, for the purpose of this document we treat these as the same entity referred to as the User.

device called a *Wallet* in which the credential/ID is stored, and may interact with a trusted *Notary* that can vouch for the contents of the ID (without sharing them with the Gatekeeper). In most real implementations, the Notary is often the same as the Wallet application on the user's device, but this need not be the case.

In principle, in these ID-based checks the Gatekeeper wishes to verify the User's age using their ID. If the Gatekeeper had some reliable way of learning only the user's age (or indeed simply the bit of whether or not the user was over the threshold age), then that is the only information the Gatekeeper needs to learn. However, during the protocol between the User and Gatekeeper, the Gatekeeper may additionally come into possession of of some *incidental data* beyond the age. This incidental data includes other info on the ID, like the User's name, but also possibly including some metadata not included as an ID attribute itself, like the issuer of the ID.

In this submission, although there are many other important aspects of ID verification and security, we primarily discuss the privacy of the User's information, and we take a model where verifying the ID's stated age can be verified via a digital signature (either the Issuer's, or the Notary's).

Discussion

Regarding this model, our research has led us to suggest the six key points stated in the abstract. We discuss each in turn:

1. Guiding Principle: Age verification methods that transmit less incidental information, to fewer entities, and to entities chosen by users, are preferable compared with other methods.

A more traditional approach to age verification is to "transmit then delete" in which the Gatekeeper (typically either the website serving age-restricted content or a third-party age verification service chosen by that website) sees some of the User's sensitive information during the process of age verification, e.g. an ID or face picture. In these non-private approaches, the User sends sensitive information to the Gatekeeper, the Gatekeeper verifies the user's age using the sensitive information, and then deletes it.

However, this forces Users to rely on the Gatekeeper to securely handle and delete their sensitive information, which is not guaranteed. This model also locks the User into an uncomfortable position where they may wish to access a specific website but not share their full ID with that specific Gatekeeper – a problem which grows in scope when observing that a single User likely interacts with many different Gatekeepers.

All else equal, avoiding transmitting that information in the first place benefits user privacy and autonomy and reduces the risk of unauthorized data access and identity theft.

We believe most legislators intend age verification laws to be functional restrictions on minors' access to age-inappropriate context and would prefer the least privacy-invasive means as possible. For example, most U.S. State age verification laws contain purpose restrictions on collected information and prohibit age verifiers from retaining or selling data collected during age verification.[2] We believe the best way to adhere to that requirement is for Gatekeepers to never receive that sensitive information at all, and for Users to choose and limit the few entities that do.

Two methods described in the next session meet this goal, making them preferable to other equally rigorous methods.

2. *Trusted Notaries* (a.k.a. selective disclosure via trusted attestation) and *cryptographic zero-knowledge proofs* should be prioritized by web standards as privacy-preserving ID-based age verification methods.[3] These two methods stand out for privacy because they both ensure that the only ID information learned by the Gatekeeper is limited to age, while still ensuring that some party verifies the validity of the ID.[4]

There are some trade-offs between the two privacy-preserving methods with regard to requirements and trust. However, both offer the attractive property of ensuring that a single User-chosen wallet application sees full ID info, compared to many Gatekeepers seeing full ID info.

We illustrate each method in Figure 1, alongside a "standard" non-private ID-based method, and describe them in more detail in Table 1.

[3]The W3C Verifiable Credentials (VC) specification [11] should encourage these by default.

[4]Depending on implementation, these methods may reveal the single bit indicating whether the User meets the age requirement, or may reveal the exact age or birth date of the User. The difference between these is important, but all these methods share the property that no *other* information on the ID is shared with the Gatekeeper.

Non-Private Baseline: Standard ID-based age verification (see Figure 1 and Table 1). The User sends a picture (or other digital representation) of their full ID to the Gatekeeper. The Gatekeeper checks the ID and outputs whether the age criterion was met. The Gatekeeper is often required by law or policy to delete or anonymize the collected ID data. This method is non-private and less secure because the User must send their full ID picture to many different Gatekeepers and trust each one to handle and/or delete it securely.

Private Method 1: Trusted Notaries (a.k.a. Selective Disclosure via Trusted Attestation)[5] (see Figure 1 and Table 1).

Before the age verification check, the Gatekeeper selects some set of Notaries it will trust. Separately, the User picks a Notary (often incorporated into a digital Wallet application) and sends their ID to the notary for verification. The Notary verifies the ID, and the User is given a digitally signed *attestation* indicating that the Notary vouches for the ID being valid and age over the threshold. During an age verification check, The Wallet sends that attestation of ID attributes to the Gatekeeper who can verify it using the Notary's public key. The Notary sees full ID information, but the Gatekeeper does not. The Notary must be trusted by the User (since it sees the ID) and the Gatekeeper (since it is relied upon for verifying the ID). In many existing implementations, the Notary is the same entity as the user's Wallet, but this need not be the case.

Existing examples of trusted notaries in the form of digital wallets include Apple Wallet [1], Dock Wallet (VCs) [3], Microsoft Entra (VCs) [8], and Colorado Digital ID [12].

Private Method 2: Cryptographic Zero-Knowledge Proof (see Figure 1 and Table 1). This works when IDs contain digital signatures that can be verified with the Issuer's known public key. While traditional ID check might have the Gatekeeper verify the ID's signature, here the user's Wallet runs ZKP *Prover* code to produce a proof that the ID has a valid signature and meets the age limit, and the Gatekeeper runs ZKP *Verifier* code to confirm the proof. This does not require the Gatekeeper to trust the Wallet – if the Wallet deviates from the protocol or tries to validate a bad ID, the ZKP Verifier code will reject it.

From a cryptographic standpoint, the Wallet and the User are the same entity. In practice, the Wallet is software on the User's device, and the User must trust it to handle the ID on the local device (similar to how the User must trust the Camera app to store a picture of an ID). Unlike the Trusted Notary method, though, the Wallet/Notary is not performing the ID verification itself, so the ID will never be sent over the network. Additionally, because the Notary does not need to be trusted by the Gatekeeper, the User could choose any Wallet that implements the protocol.[6]

Examples of existing ZKP-based implementations of ID-based age verification include Google Wallet [6], Self [10], FS'24 [4], RWG+'23 [9], ZKPassport [13], and ongoing work by these authors [7].

3. While the privacy-preserving methods described in the previous section are promising for protecting incidental ID information, **on their own they do not conceal other potentially sensitive non-ID-content information**. In particular, in naive implementations of the zero-knowledge proof approach, and sometimes in the trusted attestation approach, the Gatekeeper may learn *which issuer* issued the ID, revealing details about the User (e.g. which state/nation issued their ID). The Gatekeeper will also naturally receive other metadata like timestamps and User IP addresses, and other network information. Additional checks like "liveness checks" may be difficult to add to the system in a privacy-preserving way.

It is possible to incorporate additional checks, like freshness checks and revocation/blocklists checks, to the ZKP system [7]. Similar mechanisms could also likely be incorporated into Trusted Notary methods. However, these features are not yet present in major implementations. Handling these additional items is also important for privacy and is overlooked in current systems.

4. When doing privacy-preserving age verification **the User should be clearly shown, and must consent to, the information that will be transmitted**. They should know whether the system will

[5]The method we call "Trusted Notaries" is often called "Selective Disclosure" in the context of a digital Wallet app. However this nomenclature has two issues: First, it avoids the reality that the Gatekeeper is trusting the disclosing party to honestly vouch for the signed statement. And second, although current implementations mostly have a digital Wallet vouching for the ID, in principle the Notary and the Wallet could be two different parties. A more general term for this, that works with existing cryptography literature, is "Trusted Attestation," however this often comes with connotations of hardware security or code signed by a hardware Trusted Platform Module. To avoid these pitfalls we use the term "Trusted Notaries."

[6]For ID verification we favor proof methods that avoid the "trusted setup" needed by some ZKP systems. For more on this see the discussion on the libZK page [5].

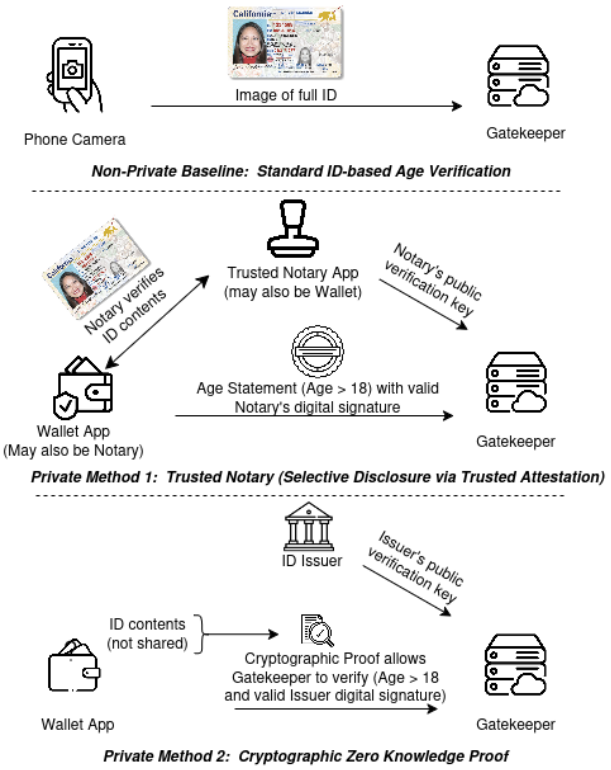

Figure 1: Illustration of privacy-preserving ID verification methods and a non-private baseline

share only their age, their age plus birth date, their age plus name, their full ID, or other information.

This property is crucial for consent-based notions of privacy. If an age verification system chooses, for any reason, to transmit age *and another attribute* (e.g. name), that fact should be made explicit to the user. We suggest adopting an approach similar to application permission dialogues that alert users when an app accesses their camera, location, or other sensitive resources.

5. In jurisdictions where laws require age verifiers to delete "identifiable information" (a provision common in many age verification laws[7]), **standards should err on the side of classifying more data as "identifiable"** since much information becomes identifying when joined with external information.

In particular, "matchable" information (including cryptographic hashes of ID information) should be treated as identifiable. A common misconception is that sharing a hash (or "fingerprint") is private because

[7]At the time of writing, 23 of 24 U.S. State age verification laws prohibit age verifiers from retaining identifiable information after access is granted.

Method	Summary	Privacy against Gatekeeper?	Gatekeeper trusts Third Party?	Primary Issues
Standard ID-based Age Verification (Non-private Baseline)	User's device acts as untrusted image storage; User transmits full ID information to Gatekeeper; Gatekeeper checks ID	No (Full ID information sent to Gatekeeper)	No (Gatekeeper correctly rejects ID if wallet compromised)	All Gatekeepers see full ID information
Trusted Notaries (Private Method 1) (e.g., Apple Wallet [1], Dock Wallet (VCs) [3], Microsoft Entra (VCs) [8], Colorado Digital ID [12])	User obtains signature from trusted Notary (may be the same entity as user's Wallet). The trusted Notary performs ID verification, a signed attestation of the age result is sent to the Gatekeeper	Yes (Gatekeeper only learns attestation of validity)	Yes (Gatekeeper relies on wallet's inspection of ID)	Malicious Notary can cause invalid ID to pass; Notaries see ID information
Cryptographic Zero-Knowledge Proofs (Private Method 2) (e.g., Google Wallet [6], Self [10], FS'24 [4], RWG+'23 [9], ZK-Passport [13], ongoing [7])	User's Wallet generates a *cryptographic zero-knowledge proof* (ZKP) by running ZKP Prover code, generating an object that proves age and validity under Issuer's public key.	Yes (Gatekeeper only learns proof of validity), although most implementations reveal Issuer	No (Gatekeeper runs ZKP Verifier code which will reject an invalid ID)	Issuer must digitally sign IDs so they can be verified with known public key; hiding Issuer requires more attention

Table 1: Table comparing privacy-preserving methods for ID verification

it cannot be inverted to recover the original data. But the same input always produces the same hash, so the hash can be used for matching across systems. By standard academic definitions, this is not private: such hashes act as linkable pseudonyms for the underlying information even if they are not reversible [2].

6. As a baseline privacy measure, in all age verification scenarios – including those that do not maximize privacy – **standards should encourage purpose restriction and prompt deletion of any incidental information collected**. This aligns with data protection and privacy best practices, even when the Gatekeeper is not legally required to limit data collection.[8]

Conclusion

In conclusion, we encourage age verification standards to prioritize methods that transmit less information over methods that rely on Gatekeepers to delete information after the fact.

We identify two promising methods for ID-based age verification: Trusted Notaries and Cryptographic Zero-Knowledge Proofs. Both methods have existing implementations. While each has its trade-offs, we strongly prefer them to other ID-based approaches because they confine sensitive ID information to a single Notary or Wallet application, rather than exposing it to many Gatekeepers. Furthermore, because the Wallet is chosen by thhe User, Users can express a preference for systems they find more trustworthy – particularly with the zero-knowledge proof approach where an interoperable protocol would enable any Wallet to be used without requiring the Gatekeeper to trust it.

In addition, we note several additional desirable properties of ID-based age verification: Protect metadata to the extent possible including Issuer identity, ensure Users know what data they are sharing and with whom, adopt a broad definition of "identifying information" to avoid de facto pseudonyms, and encourage purpose limitation and data minimization as standard best practices.

Finally, although this work focuses on privacy in ID-based age verification, we do not specifically endorse ID-based age verification in general, nor do we claim that its privacy risks are greater or lesser than those of other approaches.

These comments reflect our own individual views.

References

[1] Apple Inc. *Add your driver's license to Apple Wallet*. Accessed August 7, 2025. Mar. 2025. URL: https://support.apple.com/en-us/111803.

[2] Levent Demir et al. "The pitfalls of hashing for privacy". In: *IEEE Communications Surveys & Tutorials* 20.1 (2017), pp. 551–565.

[3] Dock Labs. *Selective Disclosure: Choose What Data To Share*. 2024. URL: https://www.dock.io/post/selective-disclosure.

[4] Matteo Frigo and abhi shelat. *Anonymous credentials from ECDSA*. Cryptology ePrint Archive, Paper 2024/2010. 2024. URL: https://eprint.iacr.org/2024/2010.

[5] Matteo Frigo and abhi shelat. *libZK: a zero-knowledge proof library*. Tech. rep. draft-google-cfrg-libzk-00. Internet Engineering Task Force, Mar. 2025. URL: https://datatracker.ietf.org/doc/draft-google-cfrg-libzk/00/.

[6] Google. *New ways to verify your age and identity with Google Wallet*. Accessed: 2025-07-14. Google. 2024. URL: https://blog.google/products/google-pay/google-wallet-age-identity-verifications/.

[7] Shuang Liu and Sarah Scheffler. *Privacy-preserving Age Verification based on Improved Verifiable Credentials Framework*. https://conpro25.ieee-security.org/papers/liu-conpro25.pdf. May 2025.

[8] Microsoft. *Introduction to Microsoft Entra Verified ID*. Accessed on 2025-08-07. June 2025. URL: https://learn.microsoft.com/en-us/entra/verified-id/decentralized-identifier-overview.

[8] Regulators should also include purpose restrictions and requirements for age verifiers to delete incidental data in new policies, however, we recognize that regulator choices are out of scope for this workshop.

[9] Michael Rosenberg et al. "zk-creds: Flexible anonymous credentials from zksnarks and existing identity infrastructure". In: *2023 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2023, pp. 790–808.

[10] Self Docs. *Self Protocol*. Last updated July 21, 2025. 2025. URL: https://docs.self.xyz/.

[11] Manu Sporny et al. *Verifiable Credentials Data Model v2.0*. Tech. rep. Accessed: 2025-07-14. World Wide Web Consortium (W3C), May 2025. URL: https://www.w3.org/TR/vc-data-model-2.0/.

[12] State of Colorado. *myColorado State of Colorado's Official Mobile App*. 2025. URL: https://mycolorado.gov/colorado-digital-id/verify.

[13] ZKPassport. *ZKPassport Introduction*. Accessed: 2025-08-07. 2025. URL: https://docs.zkpassport.id/intro.