Schema Design

Suggest Edits

Intro

Decentralized identity specifications are being actively developed by a broad range of communities and stakeholders who are committed to the vision of user-controlled, privacy-preserving identity management. There is a solid pipeline of standards-track specifications enabling building out of the decentralized identity ecosystem, with the foundational specifications reaching standard status.[^1]

While the primary specification influencing credential schemas is the VC data model, other specifications, such as those covering signature suites, as well as issuance and exchange protocols, can also impact the security, privacy, and usability of credentials for the user.

Cross-Cutting Considerations

Consider ability for relying parties to consume the credentials

While some ecosystems may seamlessly support ZKPs signature suites, this cannot always be assumed -- especially given the relatively limited library availability (and possible concerns around maturity of related specifications). For that reason, it is helpful to give subjects the option to receive credentials that are signed via schemes with broad, interoperable support like JWT signatures.

Credential issuance protocols allow subjects and issuers to reach agreement on which signature scheme to use, so this guidance doesn't imply a "downgrade" to the minimal option.

Issuers of ZKP credentials may include a higher amount of sensitive attributes in a single credential (because it expects the wallet to handle reducing disclosure), but a non-ZKP credential should minimize attributes, per theVC Data Model privacy recommendations . ZKPs impose tooling requirements on wallets to manage, which may be feasible in a closed system where one vendor supplies all the pieces, but difficult otherwise.

One alternative to ZKP-enabled selective disclosure of PII-containing credentials is supplemental VCs for authorization of remote query (e.g. API-based or oracle-based) to trusted intermediaries.

Ensure credential status registries, such as revocation lists, do not leak additional information

The VC Data Model's guidance on credential status registries can be difficult to reconcile: it mentions that revocation should not reveal additional information (about the subject, holder, the VC, or verifier) but also states that the issuers can disclose revocation reason, and that they should distinguish between revocation for cryptographic integrity vs a revocation status change.

In practice, it tends to be clearest for implementers (and safest for users) to use simple inclusion approaches to revocation or status registries, such as using theBitstring Generation Algorithm , which only reveals whether a credential index is in the set of revoked credentials, with no additional information about reason. Contents of status registries require special scrutiny, particularly since they may be on-chain. Because it can be hard to predict how any additional information can be discovered, correlated, and potentially misused, it is best to avoid providing any additional signals.

Allow reissuance and new DIDs

Similar to how some use different crypto addresses to avoid correlation, some users may prefer to use different DIDs for different credentials. In fact, use of a new DID for every credential is generally considered best practice. This doesn't manifest as a usability issue for subjects because their wallet software can manage this for them, both on the issuance and exchange (if proof of control is required) side.

In case the subject needs the credential to be issued to a different DID (key compromise or other concern), it's recommended to use the refreshService property to allow a subject's wallet to easily discover a way to request a new one. The issuer should revoke the previous and re-issue in response.

Recommendations

The following are recommendations for safe practices while making minimal assumptions.

Reduce attributes required of the credential subject through fit-for-purpose criteria

The Verifiable Credentials Data Model provides general guidance on theprinciple of data minimization , involving reducing

the amount of personal information revealed to verifiers, either through issuance of multiple single-attribute credentials or through data minimizing signature suites such as ZKPs. Lack of library availability, or concerns about maturity, may discourage some implementers from relying on the latter. At the same time, it can be difficult for issuers to envision how to property atomize credentials into single-attribute credentials: not all attributes are meant to be mixed and matched in different combinations.

A simpler approach is to tailor credential schemas to one specific use case, and avoiding use of one credential schema for multiple use cases. As an example, the Verite schema for a KYC/AML compliance credential is intended only to communicate KYC/AML compliance for a given country -- not adjacent use cases (e.g., excluding residents of states with stricter regulatory requirements). This separation enabled a KYC/AML claim with minimal subject attributes -- not requiring detailed residence information about the subject.

Prefer composability of credentials

Along with reducing the data required of a subject, careful consideration of the intended use of credentials help enable composable, stackable credentials that can be aggregated into presentations requested by relying parties. The combination of credential exchange protocols and wallet software enable this to happen without additional burden to the user.

Composing and stacking credentials is another way to achieve some effects enabled by ZKPs, while not disclosing more information needed for a given scenario.

Continuing the residency restriction example described above, suppose a verifier wanted to restrict to users that have been KYCd, and additionally, do not live in a specific state. With the composable approach, the subject would be issued a KYC credential and a separate residence credential, which would be combined to support the presentation requested by the verifier[^2].

Prefer reuse of existing schemas and vocabularies

There are not yet standard repositories of VC schemas and vocabularies, but often it's helpful to reuse any source of structured data appropriate to the domain. Schema.org is often a good starting place for the broadest expression of types and relations, and many domain-specific vocabularies are designed to fit within schema.org.

Re-use of schema.org terms is highly recommended, even if it's unclear whether it'smeant to apply to the use case. This is more apparent in practice than in documentation, but the usual practice is to re-use terms that already exist rather than coin new ones, even where this means slight changes to definition, extending the domain and range or properties.

Notes

1. W3C Verifiable Credential Data Model
2. is endorsed as a W3C recommendation, and theW3C Decentralized Identifier Data Model
3. is a proposed W3C recommendation.
4. But note that ZKPs could offer further improvements on the credential contents, more easily expressing "not a resident of" certain locations Updated5 months ago
5. Table of Contents
6.
    - Intro
7.
    - Cross-Cutting Considerations
8.
        - Consider ability for relying parties to consume the credentials
9.
        - Ensure credential status registries, such as revocation lists, do not leak additional information
10.
        - Allow reissuance and new DIDs
11.
    - Recommendations
12.
        - Reduce attributes required of the credential subject through fit-for-purpose criteria
13.
        - Prefer composability of credentials
14.
        - Prefer reuse of existing schemas and vocabularies

15.
    - Notes