ANONYOME LABS

Multiple Personas:

Enhancing Privacy in Decentralized Identity Architectures

The free to consumer nature of today's internet is built on the premise that digital content and services 'want' to be free—if advertisements, usage tracking, and personal data correlation can offset the invested costs. In the physical world, ads are creatively displayed to encourage customers to purchase the advertised products. Online, vendors can track where their ads go, who sees them, whether purchases are made, conditions related to purchases, as well as a plethora of personal information about customers, which they usually share with other vendors. All of this makes online advertising significantly more effective—but it comes at a cost.

Modern online advertising and service usage metrics retrieve so much data about users that it is sometimes referred to as surveillance capitalism. Under this new business model, users are susceptible to notable privacy risks that they often do not understand and struggle to avoid.

Everyone on earth is born in a country that can provide them with a legal identity, which is often documented with a birth certificate. This enables people to perform a variety of legal activities, such as going to school, working, banking, etc. These scenarios require a person's actual legal identity. In decentralized identity (DI) architectures, the digital

representation of the birth certificate can include its own decentralized identifier (DID) and be issued as a verifiable credential (VC).

A persona is also a digital identity representation and can be created by anyone to represent them in a variety of scenarios (e.g., home, work, shopping, social media, etc.). An easy way to describe a persona for work-related scenarios is as a business card. On business cards, people normally use a formal name, a work phone number, a work email, etc. The purpose of creating a business card with contact information is to help people be reachable during work hours—and still take weekends!

Extending the business card illustration, people can create unique contact sets for hobbies, travel, purchasing, trade shows, etc. Personas created for temporary events (e.g., selling a car, buying something from a foreign country) can safely be retired once the need for them ends.

DI architectures introduce significant security and privacy improvements, but they don't always eliminate user-specific tracking. One way to disrupt user-specific tracking is through personas. This whitepaper describes relevant portions of DI architectures and illustrates how legal identities and personas facilitate stronger privacy control.

Decentralized identifiers (DIDs)

According to the W3C Decentralized Identifiers specification, “Decentralized identifiers (DIDs) are a new type of identifier that enables verifiable, decentralized digital identity. A DID refers to any subject (e.g., a person, organization, thing, data model, abstract entity, etc.) as determined by the controller of the DID.” Further, “Each DID document can express cryptographic material, verification methods, or services, which provide a set of mechanisms enabling a DID controller to prove control of the DID.”

A DID is a URI (resembling URLs) that consists of a scheme (did:), a method identifier, and an identifier that is unique within the method’s identifier space. A DID is depicted as follows:

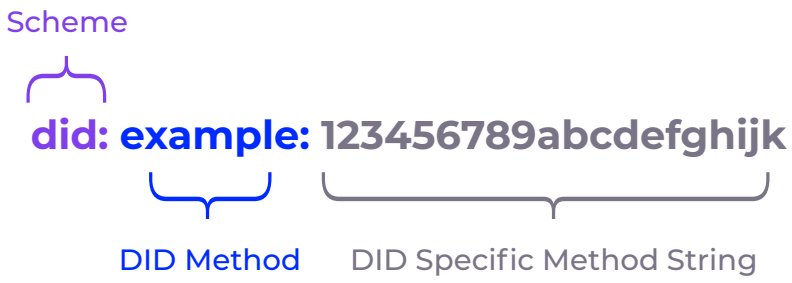


Figure 1: Elements of a DID

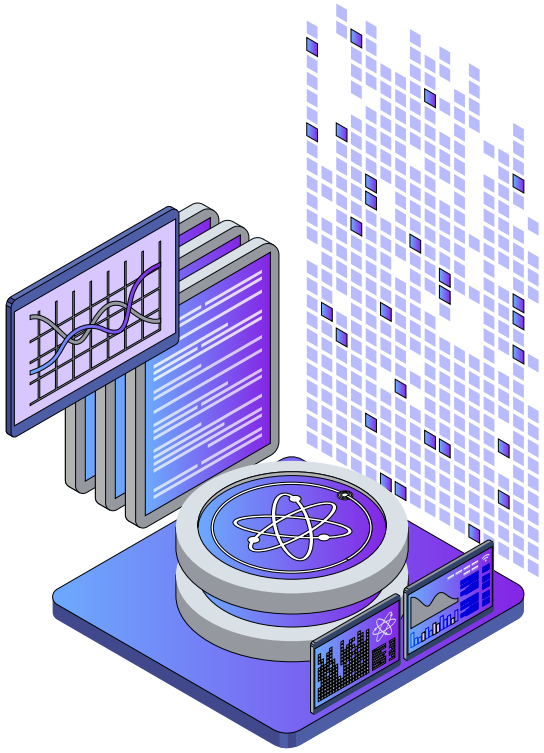
DIDs are used to reference and retrieve related DID documents (DIDDoc) similar to the following:

```
{
  "@context": [
    "https://www.w3.org/ns/did/v1",
    "https://w3id.org/security/suites/ed25519-2020/v1"
  ]
  "id": "did:example:123456789abcdefghi",
  "authentication": [{
    // used to authenticate as did:...fghi
    "id": "did:example:123456789abcdefghi#keys-1",
    "type": "Ed25519VerificationKey2020",
    "controller": "did:example:123456789abcdefghi",
    "publicKeyMultibase": "zH3C2AVvLMv6gmMNam3uVAjZpfkcJCwDwnZn6z3wXmqPV"
  }]
}
```

Figure 2 – A simple DID document from the W3C DID specification

A DID subject is the entity identified by a DID and the associated DIDDoc. By definition, a DID subject can refer to a real person (e.g., legal identity), as well as a logical representation of a real person (e.g., persona). DIDs can even refer to inanimate objects used on the Internet of Things.

Privacy and personal data



In recent years, the detriments of personal data tracking have been widely publicized. The US Government Accountability Office regularly publishes reports about how data collection poses risks to consumer privacy (see Consumer Data: Increasing Use Poses Risks to Privacy). In some market segments, there is still some confusion between data privacy and data security (see Forbes article, Data Privacy Abuse Continues Because We Struggle To Define The Problem). These make safe online interactions difficult for the average consumer.

Great care is now being taken to protect online databases and services. In many jurisdictions, legal rules are being established to define how operators must handle personal information. These pieces of legislation also define penalties for data handlers that collect, analyze, use, or resell personal data contrary to legal requirements. However, despite these rules and penalties, accidental disclosures continue to happen, and digital thieves will continue to operate contrary to established laws.

While a single data item, such as a person's name, may not itself be a unique identifier, it often becomes sufficiently unique when combined with other related data elements (e.g., birth date, home address, financial information, etc.) and that combination can often be used to analytically identify a specific individual. New and emerging analytics methods routinely uncover activities that one would normally consider to be private (e.g., purchasing books, event attendance, social interactions with online friends) and then can computationally associate them with a specific individual.

The emergence of DIDs, which are unique identifiers, has introduced numerous benefits with related core features, such as decentralized identity management, simplified cryptography operation, individually controlled identification data, end-to-end encrypted communications, cryptographic VCs, selective information disclosure, etc. These core features provide the building blocks for incalculable new and improved applications, platforms, and services.

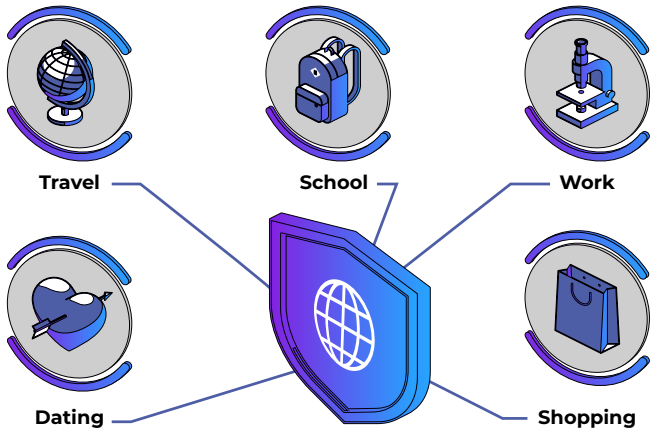
Along with the significant benefits introduced by DI methodologies, DI's determinism can also be co-opted by personal data trackers and to streamline the personal data correlation and analysis processes. For example, the presentation of a W3C VC may contain the following members: id, verifiableCredential, holder, proof, etc. By design, each of these members is usually intended to convey a holder's unique identification (for authentication purposes) to a wide range of verifiers. In some scenarios, DI's determinism may be co-opted to further enhance the data collection and analysis processes. Today's VC specifications may include privacy-enhancing features, such as zero knowledge proofs (ZKPs) but these are not universally required and may even hinder the logging requirements of legitimate verifiers (e.g., medical, government, etc.) and Trust Over IP), and more. The DI standards have matured to the point where software vendors can deliver interoperable software to the market.

Using personas to protect privacy

Modern data collection and analytics processing seeks to blur the lines between a person's various digital activities and to create a unified view of their online presence. Commonly, this is done to further targeted advertising and usage-based monetization, but it can also be performed for a wide range of motivations. To further enhance the identity analytics processes, websites that collect identity and usage statistics will often submit those to larger analytics processing services that combine many submissions in order to curate them into a variety of insights that they resell to their subscribing members. These insights not only consist of various personal data items, but also analytically discovered or inferred data points. Regardless of the motivation, the collection of one's personal information and digital activities presents a range of identity-related risks to individuals. Since the personal data collection industry is so lucrative, it is predicted to continue despite legislation seeking to regulate it.

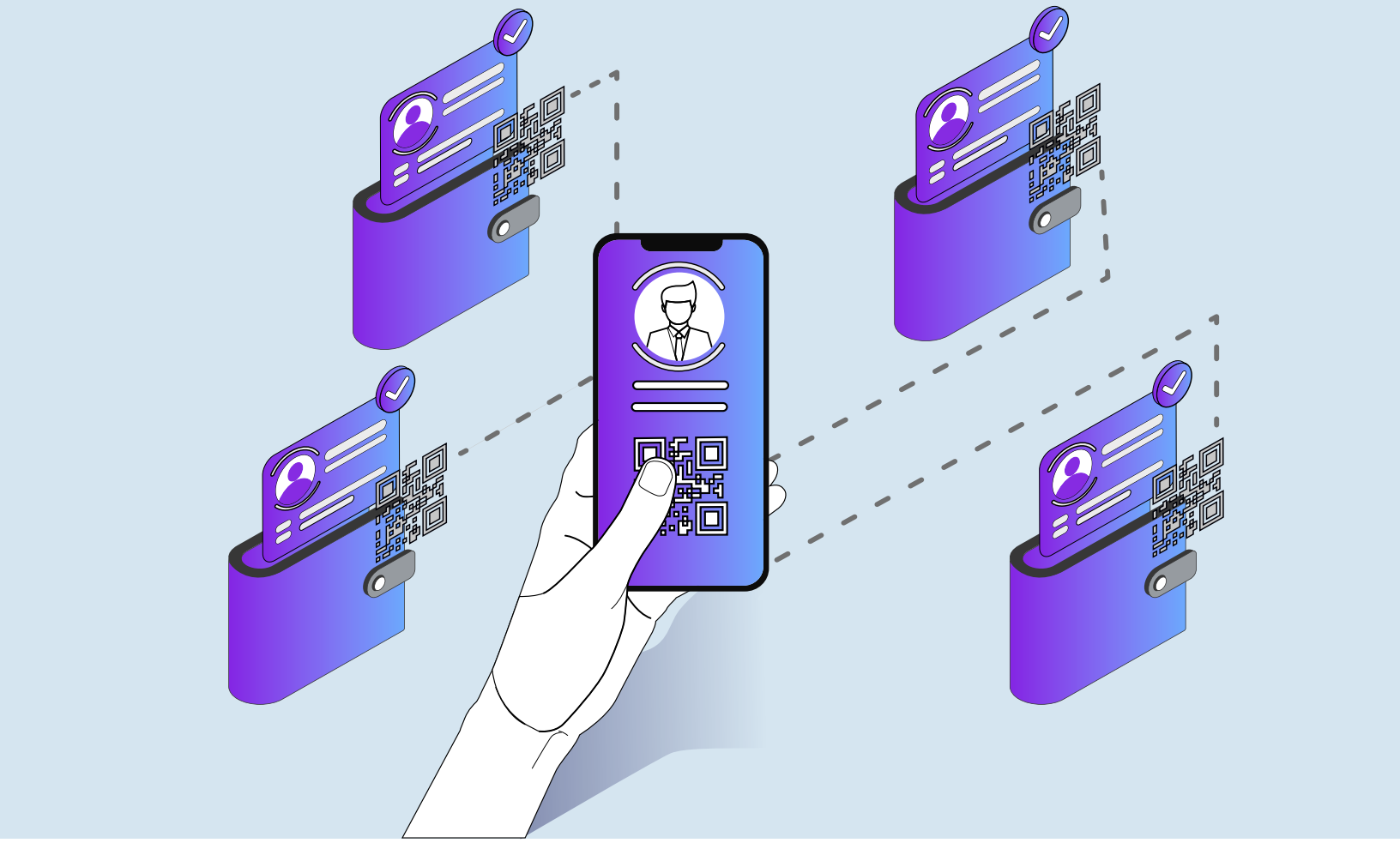
In business environments, it is customary to provide a work phone number and a work email address that are distinct from and unrelated to the personal email address and personal phone number that an individual uses in their personal non-business life. This separation of communication contact points has numerous benefits, ranging from avoiding personal distractions at work to keeping personal medical conditions or hobby-related activities separate from professional interactions. This compartmentalization of contact points enables individuals to establish and maintain their own separate work and home identities or personas. This persona-based compartmentalization process is effective, because it operates completely within the existing paradigms of how phone numbers and email addresses are currently used. Individuals can employ personas without online service operators needing to make changes or even be aware that the individual is using personas.

Just as phone numbers and email addresses have been used as unique identifiers for correlating personal data and activities, it is easy to see how DIDs may become the new correlation value of choice due to their determinism and effective immutability. Incorporating DID values into scenario-based personas will allow individuals to create and use them to help protect their privacy without regard for how DIDs and other DI values may be collected or used.



Most online scenarios do not require a person's actual or legal identity data to provide service. Rather, they typically only require that a user prove that they are authorized to access a particular account during subsequent visits. Personas fit nicely in these scenarios, because they can perform the verification authorization processes without needing to additionally provide legal or universal identifiers that can result in the unintended delivery of their full and complete personal data sets.

While distinct personas have routinely been used in work and home scenarios, they may also be used with finer granularity in additional scenarios, such as shopping, medical, social, gaming, etc. In order to help protect individual privacy from more deterministic correlation potential, using separate personas in DI environments is introduced.



Implementing personas in DI

Implementing a persona in DI is not significantly different from implementing a standard DI identity. For example, each persona will have a wallet, can manage DIDs (e.g., create, hold, use), can manage connections, can employ VCs, etc. The difference is that the app or service that manages a persona will also likely manage a user's legal identity and potentially other personas as well. This means that the implementing app will have to manage multiple DI wallets and ensure that access to them is only granted to the specific owning persona. As long as access to the wallet is limited to the owning persona, then it can operate as defined in the several DI specifications.

Verifiable credentials

On the surface, using VCs in a multi-persona context is also no different from using them in a single user context, but there are a few exceptions to the standard use cases when implementing multiple personas. The following questions will help illustrate the types of situation and new behaviors that multi-persona use cases can introduce:

1. Can one persona use another persona's VCs?

2. Can a persona use VCs owned by a legal identity?

3. Can VCs be issued to multiple personas or identities?

4. Can VCs presentations be proven using privacy-preserving methods (e.g., ZKPs)?

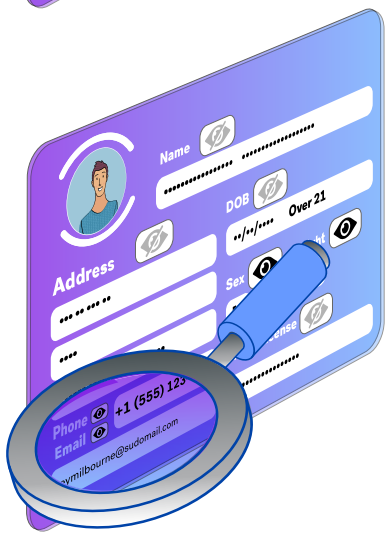
These questions will be addressed next. While the purpose of this whitepaper is to introduce the new concept of operating in multi-persona environments, it is also intended to propose how this is to be done and how it can be done within existing standards and capabilities. Adhering to the features provided in existing specifications will make the creation and adoption of multi-persona operations easier to achieve. When existing specifications do not facilitate a needed feature or capability, new material may be proposed.

AnonCreds

The AnonCreds specification was developed to enable VC usage in a privacy-preserving manner. In the most general sense, AnonCreds offers three privacy-enhancing capabilities:

1. Anonymity: Holders (e.g., credential subjects) may respond to VC proof request presentation responses without disclosing their identity. This lets holders maintain their privacy while still providing a verifiable response to a proof request.

2. Threshold responses: Holders may respond with threshold responses rather than providing specific actual personal data. For example, if a proof request asks, “Are you a local resident?”, a holder of a government-issued credential may prepare a response that essentially states “Yes” in a properly formatted manner that makes their response cryptographically verifiable without disclosing either their subject DID or actual specific information, such as their home address.



3. Partial responses: When a VC proof request solicits data elements, a holder may decide whether to provide all of the requested data or only some data elements. For example, if a verifier asks for 3 fields (e.g., name, email, phone number), a holder may choose to reply with 2 (e.g., name and email) while omitting the other.

From section 9.1 of the AnonCreds specification, a sample AnonCreds presentation request is generated by a verifier in JSON format, as follows:

```
{
    "nonce": "verifierNonce",
    "name": "pres_req_1",
    "version": "0.1",
    "requested_attributes": {
        "attr1_referent": {
            "name": "name"
        },
        "attr2_referent": {
            "name": "sex"
        },
        "attr3_referent": {
            "name": "phone"
        },
        "attr4_referent": {
            "names": ["name", "height"],
            "restrictions":
            {
                "cred_def_id":"NcYxiDXkpYi6ov5FcYDi1e:3:
   CL:NcYxiDXkpYi6ov5FcYDi1e:2:gvt:1.0:TAG_1",
                "issuer_did": "NcYxiDXkpYi6ov5FcYDi1e"
            }
        }
    },
    "attr2_referent": {
      "name": "sex"
    },
    "attr3_referent": {
      "name": "phone"
    },
    "attr4_referent": {
      "names": ["name", "height"],
      "restrictions": {
        "cred_def_id": "NcYxiDXkpYi6ov5FcYDi1e:3:
CL:NcYxiDXkpYi6ov5FcYDi1e:2:gvt:1.0:TAG_1",
        "issuer_did": "NcYxiDXkpYi6ov5FcYDi1e"
      }
    }
  },
  "requested_predicates": {
    "pred1_referent": {
      "name": "age",
      "p_type": ">=",
      "p_value": 18,
      "restrictions": [
        {
          "issuer_did": "did",
          "schema_id": "id"
        }
      ]
    }
  },
  "non_revoked": {
    "from": 1650876280,
    "to": 1682405051
  }
}
```

Figure 3: AnonCreds presentation request

In the above request, name represents the name of the request, requested_attributes denotes the credential elements being requested, and non_revoked specifies the validity status of the credential. In Section 9.2 of the AnonCreds specification, the following presentation response is provided:

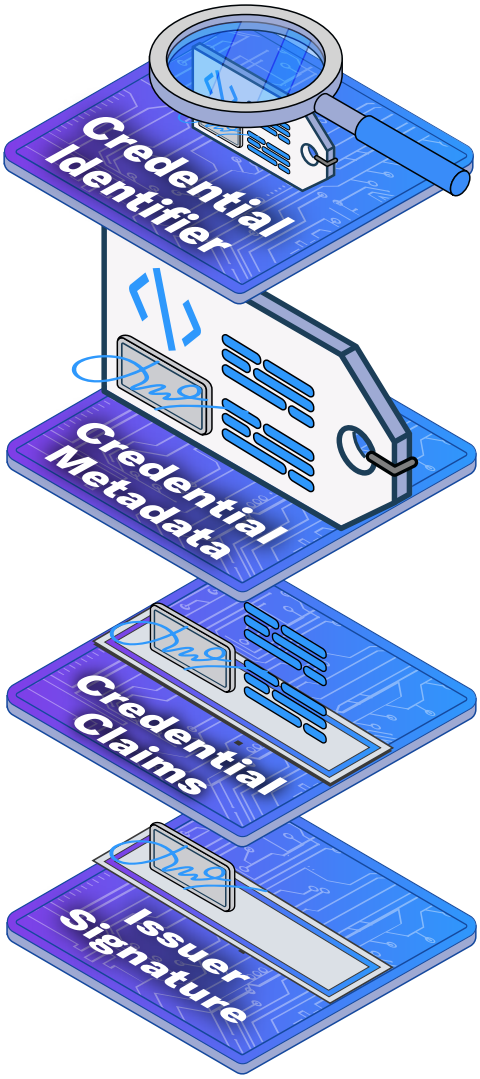
```
{
    "requested_proof": {
        "revealed_attrs": {
            "attr4_referent": {
                "sub_proof_index": 0,
                "raw": "graduated",
                "encoded": "2213454313412354"
            },
            "attr5_referent": {,
                "sub_proof_index": 0,
                "raw": "123-45-6789",
                "encoded": "3124141231422543541"
            },
            "attr3_referent": {
                "sub_proof_index": 0,
                "raw": "Bachelor of Science, Marketing",
                "encoded": "12434523576212321"
            }
        },
        "self_attested_attrs": {
            "attr1_referent": "Alice",
            "attr2_referent": "Garcia",
            "attr6_referent": "123-45-6789"
        },
     "unrevealed_attrs": {
     },
     "predicates": {
          "predicate1_referent": {
             "sub_proof_index": 0
          }
     }
    "proof" : [] //# Validity Proof, to be checked by Verifier
    "identifiers" : [ //# Identifiers of credentials that were used for
  // Presentation building
        {
            "schema_id": "transcript_schema_id",
            "cred_def_id": "123",
            "rev_reg_id": "123_123",
            "timestamp": 1550503925
        },
        {
            "schema_id": "job_certificate_schema_id",
            "cred_def_id": "456",
            "rev_reg_id": "456_456",
            "timestamp": 1550503945
        }
    ]{
{
    "requested_proof": {
        "revealed_attrs": {
            "attr4_referent": {
                "sub_proof_index": 0,
                "raw": "graduated",
                "encoded": "2213454313412354"
            },
            "attr5_referent": {,
                "sub_proof_index": 0,
```

```
            "raw": "123-45-6789",
            "encoded": "3124141231422543541"
        },
        "attr3_referent": {
            "sub_proof_index": 0,
            "raw": "Bachelor of Science, Marketing",
            "encoded": "12434523576212321"
        }
    },
    "self_attested_attrs": {
        "attr1_referent": "Alice",
        "attr2_referent": "Garcia",
        "attr6_referent": "123-45-6789"
    },
```


Figure 4 – AnonCreds presentation response

The previous AnonCreds presentation request and presentation response depict how AnonCreds handles proof requests and responses. Valid responses may provide all of the data requested, only some of the requested data, or threshold assertions of the requested data (e.g., "over 21" instead of a birthdate). Of particular note is the fact there is no subject DID specified for the credential subject (holder). The absence of this subject DID enables the holder to respond to the satisfaction of the verifier without disclosing their unique identity. After such transactions, information has been exchanged and validated in an anonymous fashion. This privacy-preserving verification process is performed through the application of Camenisch-Lysyanskaya Signature (CL-SIGNATURES).

Another major differentiation between AnonCreds and W3C credentials (described below) is that AnonCreds provides a privacy-enhancing mechanism known as a link secret. Using a link secret overcomes the problem of credential correlation introduced by credential verification processes that require a holder-specific persistent identifier. Before an AnonCreds credential is issued, the holder creates a link secret, which they save in their wallet. Next, the holder creates a blinded link secret that they submit to the issuer as part of the credential request. When an AnonCreds credential is issued, the issuer uses the blinded link secret in creating the credential. This process results in giving a verifier the ability to test that the credential being verified was issued to the individual making the credential presentation. The link secret can easily be used in multi-persona scenarios, because the real person owning the personas has access to each of their personas' identity data.



W3C credentials

The VC architecture created by the W3C introduces a different design from that of AnonCreds. Most notably, for some usage scenarios, a credential Subject's DID is included in the credential and proof. Including this value has privacy implications, because it (or a derivation) can be used as a tracking identifier.

When legal identity and personas are known in advance

Section 4.4 of the W3C Verifiable Credential Specification demonstrates that VCs can be issued to multiple credentialSubjects (see example 7). This is also reiterated in Section A.1. Example 7 demonstrates that a single credential could be issued to two spouses and that either spouse could use that credential for its designated purpose.

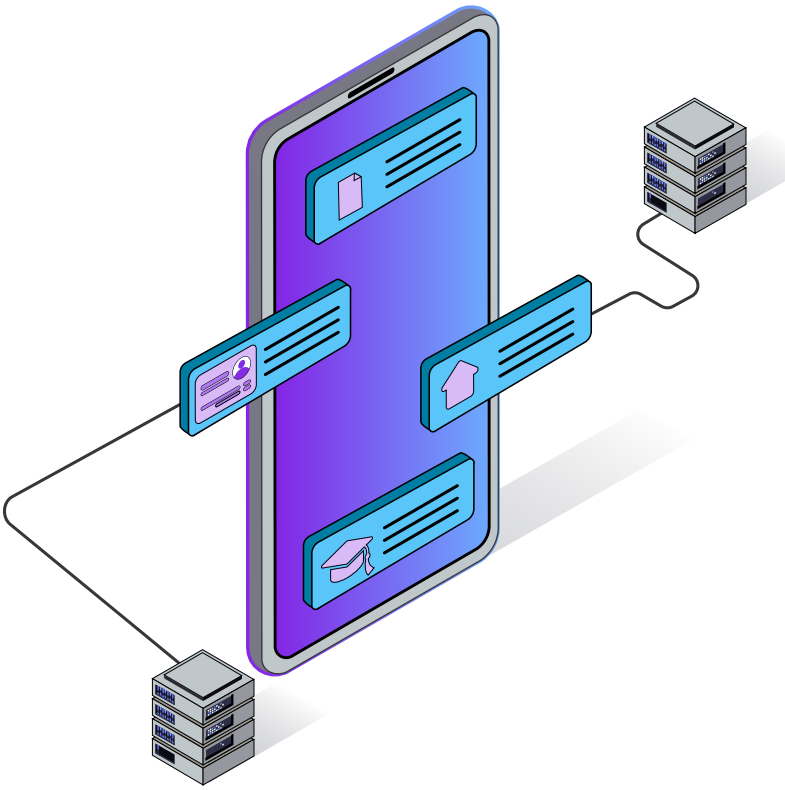
Using this method as a model, an issuer could issue a credential to a user's legal identity and one or more of their personas. This would enable either of the identities to use that credential within the scope of their purpose. This method would enable multiple personas to use a shared credential, but it would require the personas to be enumerated in advance of the credential's issuance.


```
{
  "@context": [
    "https://www.w3.org/2018/credentials/v1",
    "https://www.w3.org/2018/credentials/examples/v1"
  ],
  "id": "http://example.edu/credentials/3732",
  "type": ["VerifiableCredential", "RelationshipCredential"],
  "issuer": "https://example.com/issuer/123",
  "validFrom": "2010-01-01T00:00:00Z",
  "credentialSubject": [{
    "id": "did:example:ebfeb1f712ebc6f1c276e12ec21",
    "name": "Jayden Doe",
    "spouse": "did:example:c276e12ec21ebfeb1f712ebc6f1"
  }, {
    "id": "did:example:c276e12ec21ebfeb1f712ebc6f1",
    "name": "Morgan Doe",
    "spouse": "did:example:ebfeb1f712ebc6f1c276e12ec21"
  }]
}
```


Figure 5 – Using multiple credentialSubjects in a W3C verifiable credential

There are pros and cons to using this method. First, this method is directly supported by the W3C Verifiable Credential Specification. Using canonical methods helps ensure that the necessary processes will be implemented by issuers, verifiers, holders, and their software providers. The downside of this method is that the various credentialSubjects could be correlated as a result of this process. For some persona implementations this is acceptable, while for others it may not be. Individual user circumstances should be considered before relying on this approach.

Masking the credential subject



The W3C specification allows for private credential proofs that mask the credentialSubject similar to how AnonCreds is structured. To accomplish this, the VCs need to be created and issued with ZKPs specified. According to the W3C Verifiable Credential Specification, holders can:

- "Combine multiple Verifiable Credentials from multiple Issuers into a single Verifiable Presentation without revealing Verifiable Credential or subject identifiers to the Verifier."
- "Selectively disclose the claims in a Verifiable Credential to a Verifier without requiring the issuance of multiple atomic Verifiable Credentials."
- "Produce a derived Verifiable Credential that is formatted according to the verifier's data schema instead of the Issuer's, without needing to involve the Issuer after Verifiable Credential issuance."

Enabling holders to define proof schemas that selectively include elements from different VCs and then to verify the included elements allows holders and verifiers to implement trust models that are very privacy preserving. This is a huge win for standardizing privacy-enhancing VC paradigms. W3C VCs implement this functionality using the same algorithm as AnonCreds, namely the Camenisch-Lysyanskaya Signature (CL-SIGNATURES), which makes them algorithmically compatible. As of W3C Verifiable Credentials Data Model v2.0, the normative statements regarding ZKPs have been expanded to allow for other ZKP schemes, such as BBS+. However, the relevant proofs associated with any new schemes also need to be added to the credential when it is created.

Despite the notable benefits, there are some similarly notable drawbacks to how W3C VCs integrate ZKP functionality. For example, W3C ZKPs currently require that CLSignature2019 signatures are specified when the credential is issued. Below is an example of a W3C VC that supports ZKPs:

```
{
  "@context": [
    "https://www.w3.org/2018/credentials/v1",
    "https://www.w3.org/2018/credentials/examples/v1"
  ],
  "type": ["VerifiableCredential", "UniversityDegreeCredential"],
  "credentialSchema": {
    "id": "did:example:cdf:35LB7w9ueWbagPL94T9bMLtyXDj9pX5o",
    "type": "did:example:schema:22KpkXgecryx9k7N6XN1QoN3gXwBkSU8SfyyYQG"
  },
  "issuer": "did:example:Wz4eUg7SetGfaUVCn8U9d62oDYrUJLuUtcy619",
  "credentialSubject": {
    "givenName": "Jane",
    "familyName": "Doe",
    "degree": {
      "type": "BachelorDegree",
      "name": "Bachelor of Science and Arts",
      "college": "College of Engineering"
    }
  },
  "proof": {
    "type": "CLSignature2019",
    "issuerData": "5NQ4TgzNfSQxoLzf2d5AV3JNiCdMaTgm...BXiX5UggB381QU7ZCgqWivUmy4D",
    "attributes": "pPYmqDvwwWBDPNykXVrBtKdsJDeZUGFA...tTERiLqsZ5oxCoCSodPQaggkDJy",
    "signature": "8eGWSiTiWtEA8WnBwX4T259STpxpRKuk...kpFnikqqSP3GMW7mVxC4chxFhVs",
    "signatureCorrectnessProof": "SNQbW3u1QV5q89qhxA1xyVqFa6jCrKwv...dsRypyuGGK3RhhBUvH1tPEL8orH"
  }
}
```

Figure 6 – W3C verifiable credential with CL signatures

The main drawback for W3C VC implementers is that privacy-preserving ZKPs are not the default credential algorithm, which means that they must be overtly selected when the credentials are designed. Additionally, if a W3C credential is not issued with a CLSignature2019 (or BBS+) signature block, then it cannot be used in ZKP processing. Since one issuer may choose to support ZKP functionality while another does not, users must understand how each W3C credential is designed before they can know whether they can or should use it in a particular scenario. If users cannot use W3C credentials without intentional effort, then they will likely ignore this feature and presume that they are not privacy preserving.

W3C credentials: identity or privacy?

At present, the W3C VC specifications appear to favor identity authentication over privacy, but as a ¬non-normative addition to the W3C VC data model, Section 7. Privacy Considerations goes into great detail about privacy and several issues that it encompasses. In Section 7.1 Spectrum of Privacy, the following figure (Figure 12 from the W3C VC data model) illustrates how credentials can span the range from highly correlatable to pseudonyms:

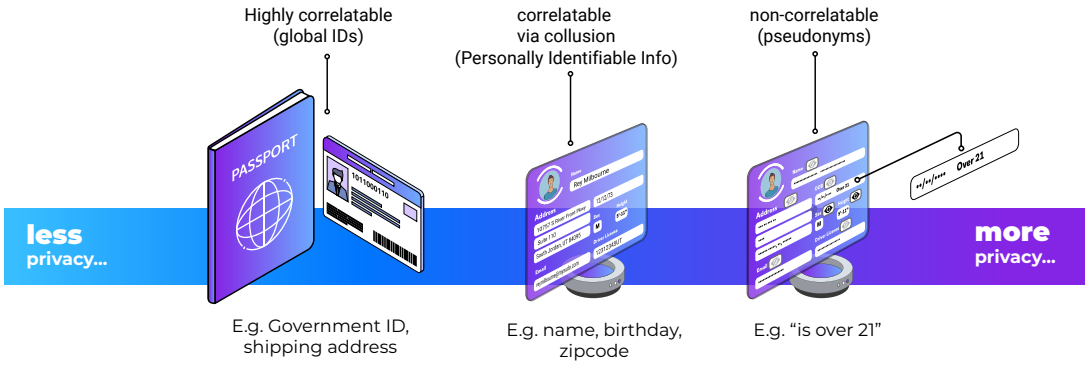


Figure 7 – How credentials span the range from highly correlatable to pseudonyms

Taking the privacy discussion further, Section 7.8 The Principle of Data Minimization describes how information disclosure best practices are migrating towards "limit[ing] the information requested, and received, to the absolute minimum necessary." It further states that, "This data minimization approach is required by regulation in multiple jurisdictions, including the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union."

Given that very far-reaching regulations, such as HIPAA and GDPR, are emerging that mandate privacy-first principles and limited data disclosure, retention, and processing policies, perhaps the W3C's VCs specifications should similarly emphasize privacy-oriented credentials as a default over those simply providing identity authentication capabilities.

If the W3C specifications made privacy-enhancing credentials the default, then both goals (identity and privacy) could easily be attained. For example, if a VC was issued using ZKP methods, it could provide the privacy necessitated by HIPAA and GDPR (e.g., "Are you over 21?"). Further, in situations where specific identity or other attributes were mandated or required, a ZKP-issued proof could also respond to a verifier's proof request soliciting specific attributes (e.g., a government-issued ID number, birthdate, etc.).

Given that ZKP-based proofs can provide both privacy-enhancing responses, as well as specific data attribute requests, it is proposed that ZKP-based proofs should become the default unless identity-only credentials are required.

Integrating existing VCs architectures

Multi-persona architectures increase privacy by segmenting what would ordinarily be an individual user's activities into a variety of persona-specific activities. A user adopting a multi-persona usage scenario will typically have a legal identity which represents their real self in addition to one or more personas.

AnonCreds is privacy preserving by design and easily integrates with the multi-persona architecture. W3C credentials also integrate easily when they are created to support ZKP signatures, but since ZKP signatures are optional features that are not added to W3C credentials by default, there is some uncertainty about whether a persona can assert a credential held by a legal identity without correlating the two identities. This is not the case with AnonCreds, which keep the correlation private. This distinction makes AnonCreds usable in all cases where privacy protections are required, while W3C credentials require some additional constraints.

Conclusion

The risks to personal privacy are notable, well-documented, and are driving nations to create privacy regulations that stipulate how (conscientious) service providers must operate and what penalties will be imposed if they don't. DI architectures have introduced dramatic advancements in [applied] cryptography, blinded interaction mechanisms, and methods for protecting personal data. However, these improvements do not anticipate new types of personal data collection, attacks on personal data stored on cloud systems, or AI-based correlation techniques. Employing methods for keeping personal data or activity-based information appropriately separated is a task that has so far been left up to technically savvy and privacy-focused users. Requiring such expertise to keep personal data private and out of the reach of enterprising data mining processes leaves typical users at the mercy and hoped for benevolence of data aggregators who financially depend on continuously enhancing data aggregation. Despite the increasing security protections, enhanced privacy protections cannot be left only to those with significant expertise and attention spans.

In order to protect against current and emerging threats against personal private information, Anonyome Labs has introduced assistive privacy protections that are easy for all users to employ. Enabling users to easily create and select situation-specific personas gives users the ability to perform activity-based segmentation of personal data without requiring the notable expertise that is often required for safe internet usage today. Combining persona architectures with DI's security improvements provides users with personally controlled privacy protections that give them an advantage over users who do not employ persona-based data and activity protections. In the same way that business cards present a business persona in work scenarios, using multiple personas in DI environments helps users increase their personal privacy … and also take weekends!