EN's Personal Identity Concept

Zeljko Milinovic, Certified Blockchain Architect, Enduring Net Tech Advisor

Ser-Huang Poon, Professor at the University of Manchester, Enduring Net Trustee

This version: November 27, 2019

This document explains how privacy of modern slavery victims and survivors can be preserved under General Data Protection Regulation (GDPR), while at the same time protecting the safety and identity of the support personnel and data holders. For the previous, we rely on Hyperledger Indy Decentralized Identifiers (DIDs) together with ZKP, and for the latter, we use a Blind Librarian Referencing System (BLRS). Enduring Net will work on developing EN Indy pools architecture.

Hyperledger Indy and DID

Hyperledger Indy is a distributed ledger purpose-built for decentralized identity leveraging blockchain technology to enable privacy-preserving digital identity. It provides a decentralized platform for issuing, storing, and verifying credentials that are transferable, private, and secure between two entities (e.g. an organisation requesting information and the holder of the private data) in a private relationship that is known only to these two entities. The holder of the data can have many such 1-1 private relationships with many entities and *vice versa*. More importantly, the data holder manages its own DIDs and makes the decision whether or not to permit an organisation to query certain attributes of the data.

The DID Model is standardized by W3C (https://www.w3.org/TR/did-core/). Specifically, a DID is a 1-1 secret relationship that is managed using public-private key and cryptographic algorithms. A DID is organically designed to incorporate privacy by design according to GDPR.

Each DID relationship is recorded on a DID document that contains the public key of the data holder. As the Sovrin Ledger provides cryptographic verification of the DID document, users can rely on Sovrin to return the public key for any DID. Such an arrangement removes the reliance on a centralized authority in existing PKI (public key infrastructure) who automatically posses the information about which key belongs to which data holder. The public-private keys can be used the other way round to give the data holder confidence that the holder is communicating with the right organisation. We have avoided personalising the data holder and the information seeker as they can both be non-human data storage devices/algorithms.

Selective Disclosure as data minimization (i.e. collection of PII that is directly relevant and necessary to accomplish a specific purpose) and *Privacy by Design* (according to GDPR) are both accomplished via DID protocols.

In summary, some of the main Indy benefits to the Enduring Net PoC:

1. Control — data holder has full control over the PII
2. Access — Access to their own data
3. Transparency — The two entities are visible to each other
4. Interoperability — with the holder's permission, the particular PII can be used by anyone on the network
5. Consent — Agreement on using their own identity
6. Existence — Allow PII owner an independent existence
7. Longevity — Exists as long as the user wishes
8. Portability — Transportable Identifiers

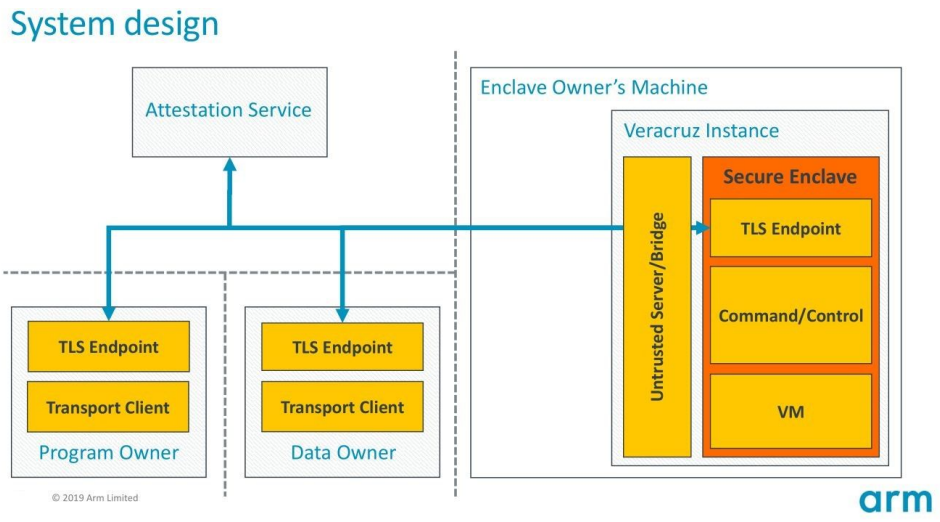
9. Minimization — Disclosure of claims are minimized
10. Protection — The rights of user's are protected.

Blind Librarian Referencing System (BLRS)

The objectives of BLRS are to protect (i) data owner, (ii) data holder, and (iii) the information seeker using a distributed support system design. It can be seen as a system where the 1-1 DID relationships reside in parallel to Enduring Net Communication Organiser ENv1. In our example below, Hannah, a social worker, is the information seeker. Alice is the data owner, and the data holders may include Alice's GP, some human rights lawyer, and police forces who keep biometric data of Alice. We assume all data holders are DSH (Data Safe Haven) with the right protection set up to ensure integrity and confidentiality of the data.

- Hannah, a user on the BLRS, sends the query "Do these (fractions of) fingerprints belong to Alice?" to e.g. a network of 100 DSHs.
- The BLRS will randomise Hannah's query with 19 other faked queries.
- The 100 DSHs will reply to provide the suggested ID of the set of 20 fingerprints, or reply to say that they have no information.
- The system replies to Hannah that e.g. 27% of the DHSs confirmed the prints belong to Alice, and provides the organisation types of the respondents (e.g. 20% law enforcement units, 7% of respondents are NGOs).
- The DSHs do not know who sent the query (i.e. Hannah in this case), nor exactly whose fingerprints are in question (i.e. Alice in this case) except that the target must be 1 of the 20 fingerprints owners. No one knows which 27 of the 100 nodes hold Alice's fingerprints.

A possible environment to mount this BLRS service is on ARM's *Veracruz* Enclave which is still under development. As shown in the figure below, the programme owner, data owner and enclave owner's machine are all protected against each other with TLS (Transport Layer Security) end points.



Identity Correlation

Identity Correlation results in a "linking" of two Digital Identities to a Digital Subject (or principal). Identity Correlation is a privacy consideration and can only be performed by an Identity Broker. With

appropriate design and the concept of ZKP (zero knowledge proof),[1] it might be possible for such an operation to be secure and fully complied with GDPR.

Under Hyperledger Indy we have a different DID for each relationship, hence identity correlation cannot be conducted automatically. It is theoretically possible, however, for two or more digital identities to be correlated via the BLRS. It might be possible that one can derive that the identity of two victims in two separate incidents actually belong to the same natural person without revealing who the person is. This area is left for future research and development.

Note: We are currently investigating MIT Enigma (2015), Hyperledger Sawtooth and other solutions provided by Intel and Samsung. Hyperledger Aries Project is also in focus of our further development.

Reference:

Derek Miller, Dominic Mulligan, Hugo Vincent, Shale Xiong (2019) "Veracruz: Privacy-preserving compute", ARM presentation at the Newton Institute VeTSS Worksop on Verified Software.

[1] ZKP (Zero-knowledge proof) primitives enable the data holder to prove to the verifier that the data satisfies a certain set of properties (knowledge) without revealing the actual data. A simple example is Apple-Pay who can authenticate payments while the cashier has zero-knowledge of the payer.