# Privacy-Preserving Claims Exchange Networks for Virtual Asset Service Providers


Thomas Hardjono
*MIT Connection Science*
*Massachusetts Institute of Technology*
Cambridge, MA 02139
hardjono@mit.edu

Alexander Lipton
*MIT Connection Science*
*Massachusetts Institute of Technology*
Cambridge, MA 02139
alexlipt@mit.edu

Alex Pentland
*MIT Connection Science & Media Lab*
*Massachusetts Institute of Technology*
Cambridge, MA 02139
pentland@mit.edu



***Abstract*—In order for VASPs to fulfill the regulatory requirements from the FATF and the Travel Rule, VASPs need access to truthful information regarding originators, beneficiaries and other VASPs involved in a virtual asset transfer instance. Additionally, in seeking data regarding subjects (individuals or organizations) VASPs are faced with privacy regulations such as the GDPR and CCPA. In this paper we a propose privacy-preserving claims issuance model that carries indicators of the provenance of the data and the algorithms used to derive the claim or assertion. This allows VASPs to obtain originator and beneficiary information without necessarily having access to the private data about these entities. We also propose a consortium trust network arrangement for VASPs to exchange signed claims about subjects and their public-key information or certificate.**



***Index Terms*—virtual assets, blockchain technology, cryptocurrency, trust network, cryptography.**


## I. INTRODUCTION

Virtual asset service providers (VASP) face a data problem. More specifically, in order for VASPs to fulfill the regulatory requirements from the FATF and the Travel Rule, VASPs need access to truthful information regarding originators, beneficiaries and other VASPs involved in a virtual asset transfer instance. However, getting access to data or information – regarding individuals and institutions involved in the asset transfer – means that VASPs must also address the challenges pertaining to data privacy and privacy-related regulations such as the GDPR [1] and CCPA [2]. On top of these issues, in the past few years there has been decreasing trust of consumers in institutions. Negative reports regarding incidents of attacks on crypto-exchanges (e.g. [3], [4]) compound this diminishing consumer trust.

We summarize these challenges as follows:

- *The Travel Rule for virtual assets*: The FATF Recommendation 15 [5] requires VASPs to retain information regarding the originator and beneficiaries of virtual asset transfers. The implication here is that cryptocurrency exchanges and related VASPs must be able to share the originator and beneficiary information for virtual assets transactions.
- *FinCEN compliance requirements*: The FinCEN rules for anti-money laundering (AML) from 2014 [6] requires that customer due diligence (CDD) be performed and includes a new regulatory requirement to identify beneficial owners of legal entity customers.
- *Decreasing trust of consumers in institutions*: Over the last decade there has been a continuing decline in trust on the part of individuals with regards to the handling and fair use of personal data [7], [8]. Pew Research reported that 91 percent of Americans agree or strongly agree that consumers have lost control over how personal data is collected and used, while 80 percent who use social networking sites are concerned about third parties accessing their shared data [9]. This situation has also been compounded by the various recent reports of attacks and theft of data (e.g. Anthem [10], Equifax [11]).

In order to begin addressing these challenges, we believe that alternative approaches to claims issuance based *decentralized data repositories* need to be explored which preservers the *privacy of the subject* (individual) and which operate based on *meaningful consent* as defined by the GDPR [1].

## II. VIRTUAL ASSETS AND VASPS

The *Financial Action Task Force* (FATF) is an inter-governmental body established in 1989 by the ministers of its member countries or jurisdictions [12]. The objectives of the FATF are to set standards and promote effective implementation of legal, regulatory and operational measures for combating money laundering, terrorist financing and other related threats to the integrity of the international financial system. The FATF is a "policy-making body" which works to generate the necessary political will to bring about national legislative and regulatory reforms in these areas.

With the emergence of blockchain technologies, virtual assets and cryptocurrencies, the FATF recognized the need to adequately mitigate the money laundering (ML) and terrorist financing (TF) risks associated with virtual asset activities. In its most recent Recommendation 15 [5], the FATF defines the following:

- *Virtual Asset*: A virtual asset is a digital representation of value that can be digitally traded, or transferred, and can be used for payment or investment purposes. Virtual assets do not include digital representations of fiat currencies, securities and other financial assets that are already covered elsewhere in the FATF Recommendations.

- *Virtual Asset Service Providers* (VASP): Virtual asset service provider means any natural or legal person who is not covered elsewhere under the Recommendations, and as a business conducts one or more of the following activities or operations for or on behalf of another natural or legal person: (i) exchange between virtual assets and fiat currencies; (ii) exchange between one or more forms of virtual assets; (iii) transfer of virtual assets; (iv) safekeeping and/or administration of virtual assets or instruments enabling control over virtual assets; and (v) participation in and provision of financial services related to an issuer's offer and/or sale of a virtual asset.

In this context of virtual assets, transfer means to conduct a transaction on behalf of another natural or legal person that moves a virtual asset from one virtual asset address or account to another. Furthermore, to manage and mitigate the risks emerging from virtual assets, the Recommendations states that countries should ensure that VASPs are regulated for AML/CFT purposes, and licensed or registered and subject to effective systems for monitoring and ensuring compliance with the relevant measures called for in the FATF Recommendations.

## III. The Travel Rule and Customer Due Diligence

One of the key aspects of the FATF Recommendation 15 is the need for VASPs to retain information regarding the originator and beneficiaries of virtual asset transfers. The implication of note [13] is that cryptocurrency exchanges and related VASPs must be able to share the originator and beneficiary information for virtual assets transactions. This process – also known as the *Travel Rule* – originates from under the US Bank Secrecy Act (BSA - 31 USC 5311 - 5330), which mandates that financial institutions deliver certain types of information to the next financial institution when a funds transmittal event involves more that one financial institution. This rule became effective in May 1996 and was issued by the Treasury Department's Financial Crimes Enforcement Network (FinCEN). This rule was issued by FinCEN concurrently with the new BSA record keeping rules for funds transfers and transmittals of funds.

Given that today a virtual asset on blockchain is controlled through the public-private keys bound to that asset, we believe there are other information (in addition to the customer and account information) that a VASP needs to retain in order to satisfy the travel rule [14], [15]:

- *Key ownership information*: This is information pertaining to the legal ownership of cryptographic public-private keys. When a customer (e.g. originator) presents their public key to the VASP for the first time, there must be a "chain of provenance" evidence regarding the customer's public-private keys which assures that the customer is the true owner. Proof of possession of the private key (e.g. using a challenge-response protocol, such as [16]) does not prove legal ownership of the public-private key.
- *Key operator information*: This is information or evidence pertaining to the legal custody by a VASP of a customer's public-private keys. This information is relevant for a VASP which adopts a key-custody business model in which the VASP holds and operates the customer's public-private keys to perform transaction on behalf of the customer.

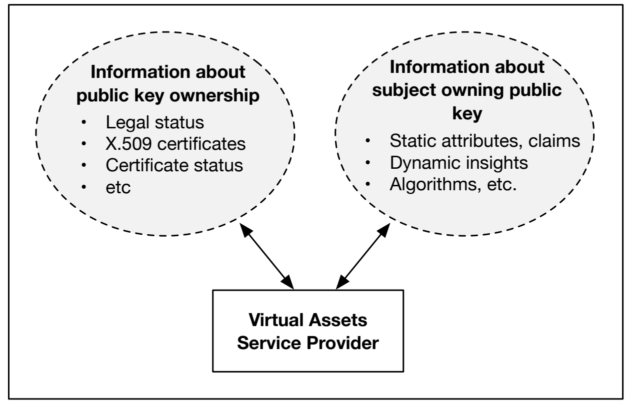


Fig. 1. The dual problem of customer CDD and customer public keys

Over the past several years financial authorities in the U.S. and in other countries have sought to modernize and more broadly enforce the anti-money laundering (AML) and terrorism financing regulations. In 2014 the US Financial Crimes Enforcement Network (FinCEN) proposed a number of Know Your Customer (KYC) requirements under "...the Bank Secrecy Act to clarify and strengthen customer due diligence requirements for: banks; brokers or dealers in securities; mutual funds; and futures commission merchants and introducing brokers in commodities" [6]. The proposed rules contained explicit customer due diligence (CDD) requirements and included a new regulatory requirement to identify "beneficial owners" of customers who are legal entities.

For FinCEN, the key elements of CDD include: (i) identifying and verifying the identity of customers; (ii) identifying and verifying the identity of beneficial owners of legal entity customers (i.e., the natural persons who own or control legal entities); (iii) understanding the nature and purpose of customer relationships; and (iv) conducting ongoing monitoring to maintain and update customer information and to identify and report suspicious transactions. Collectively, these elements comprise the minimum standard of CDD, which FinCEN believes is fundamental to an effective AML program [6].

In the past KYC verifications was considered sufficient for a financial institution when it sought to onboard a prospective customer. In recent years, however, this has evolved into broader Customer Due Diligence (CDD) programs which must be carried out throughout the relationship with the customer or client. An AML program therefore must not only include performing CDD of new customers during on-boarding, but also carrying out CDD on an on-going basis throughout the business relationship with the customer.

For financial institutions which carry-out large numbers of transactions on a daily basis (e.g. banks and investment firms), the CDD provides an on-going *assurance framework*

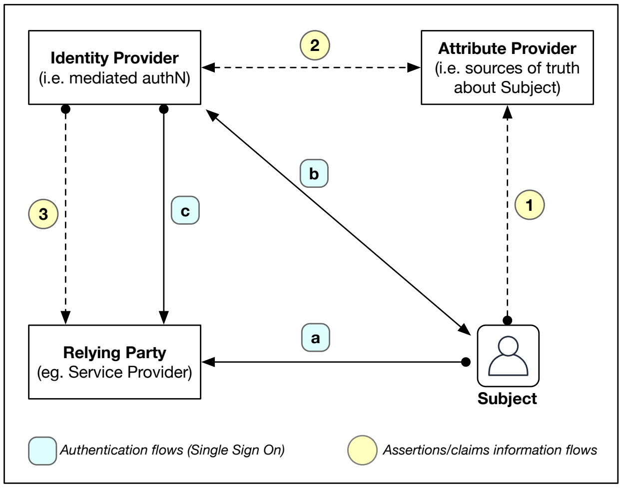


Fig. 2. The SAML Mediated Authentication & Attributes flow [18]

internally. The CDD checks includes, among others, the transaction volumes, amounts and geographical locations of the transactions (e.g. origin and destination). Using new software and networks monitoring technologies, the movements of funds can be monitored closely and suspicious activities be identified. Accurate and time data, therefore, represents a core need for these monitoring software and networks to operate correctly.

The FATF definition of virtual assets ("a digital representation of value that can be digitally traded, or transferred, and can be used for payment or investment purposes") means that VASPs – like traditional financial institutions – need to establish an effective AML program in the sense of FinCEN [6]. Furthermore, because the blockchain network as a new medium of transactions operate on the basis of public key cryptography, VASPs must additionally obtain and retain the originator/beneficiary cryptographic key ownership information as part-and-parcel of monitoring the the movements of funds. This dual problem of customer information and customer public key information [15] is illustrated in Figure 1.

## IV. Customer Due Diligence & the Identity Claims Model

The problem of customer identification, on-boarding and due diligence is not unique to VASPs, and has been a challenge for Internet service providers generally (i.e. online merchants) since the late 1990s. The promise of the Internet-based services (versus traditional brick-and-mortar shops) was that of an increase in transaction efficiency, lower costs and better convenience for the user. However, as the past two decades of Internet services has shown, the problem of consumer identification and authentication is not trivial and is closely related to the problem of authorized (consent-based) access to personal data [17].

### *A. Attributes in the SAML2.0 Model*

Online services today employ *Identity Providers* (IdP) as means to provide mediated authentication of the user (subject) on behalf of the online Service Providers (SP), such as online merchants [18]–[20]. The Service Providers are reliant on the authentication-event outcome of the IdP, and as such they are referred to also as the Relying Party (RP). The typical consumer-facing IdP issues an identifier (e.g. email address) and manages the credentials of the user (e.g. change password). When the user seeks to access services offered by the Service Provider, the user is temporarily redirected by the SP to the IdP for authentication. If the authentication is successful, the IdP issues an authentication-token (e.g. SAML2.0 tokens, Kerberos tickets) which can then be validated by the Service Provider. The IdP and the Service Provider typically have a business relationship that provides the foundation of trust between them [21].

Figure 2 illustrates the basic mediated authentication flows through the IdP in steps (a)–(c). After the IdP provides the Service Provider (Relying Party) with evidence of successful authentication in Step (c), the Service Provider now requires factual information or *attributes* (claims) about the user (subject). Here, one approach defined by the SAML2.0 specifications [18] is for the IdP to inquire to a special entity called the *Attribute Provider* (AtP) to furnish the IdP with attribute assertions or claims about the subject. In other literature (e.g. [22]), the Attribute Provider is also known as the *Claims Provider* (CP). Thus, in Step (1) of Figure 2 the subject provides consent or authorization for the Attribute Provider to release information to the IdP in Step (2). The IdP forwards the assertions or claims in Step (3) to the Service Provider. Alternative flows are possible, such as when the signed claims are delivered to the SP through the Subject.

It is worth noting that the main business reason the IdP mediates the flows of Figure 2 is because the IdP provides a scaling factor in the face of multiple Service Providers and multiple Attribute Providers. The user needs only one (or few) credential account at the IdP in order to engage multiple Service Providers (e.g. online merchants). Today, the dominant IdP function is provided mainly by social media platforms [17].

From the VASP perspective, the flows in Figure 2 provides the rudimentary mechanism for a VASP to obtain customer information in the form of signed claims. Thus, the VASP is the relying party because it is reliant on the Attribute Provider (Claims Provider) to furnish it with information about the subject seeking the services of the VASP (e.g. subject request transfer of virtual assets).

### *B. Authorization to Access Protected Claims in UMA2.0*

While the IdP model [18] provides a solution for the needs of Web Single Sign On (Web-SSO) starting in the late 1990s, the IdP model does not solve the actal core need of Service Providers (such as VASPs), namely that of *accurate information about subjects based on data of known provenance*. Information resources (e.g. accounts information, identifiers, files, data) about users on the Internet are typically distributed across multiple IdPs, Service Providers and Claims Providers. Thus, in such a situation, the user is now faced the additional

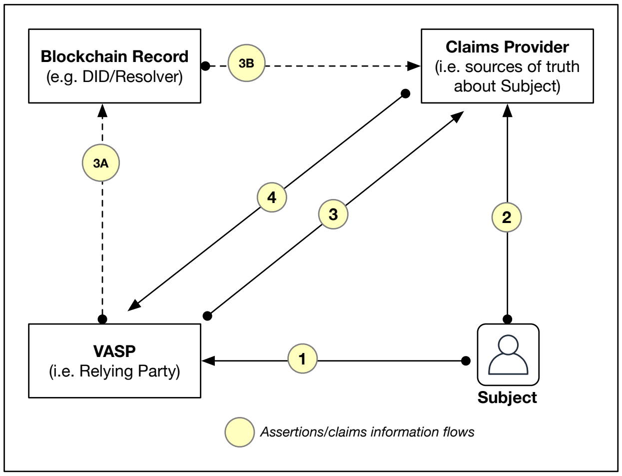


Fig. 3. The Claims Provider flow [23], [26]

problem of managing authorizations for other parties to access their multiple resources distributed on the Internet. Thus, users need a singular coherent and consistent access control policy-setting mechanism whose enforcement could be carried out across different resource points (policy enforcement points) on the Internet.

To this end the *User Managed Access* (UMA) architecture and authorization protocol [23], [24] was developed as means to standardize user-centric authorization for access to the user's information resources spread across distributed locations. An important contribution of UMA is the standardization of a protocol that realized *consent-management* in the sense of the GDPR [1]. The UMA flows complements other flows pertaining to consent-receipt management [25]. The UMA protocol employed the standard Auth2.0 authorization token [26] as the conveyance format for the consent to access the user's protected resources.

In the context VASPs, the claims belonging to a subject (user) represents a protected resource whose release to the VASP must be under the consent of the subject. As such, the UMA protocol provides a basis for VASPs to implement consent management and tracking using standardized APIs and service endpoints.

### C. Linking Claims to Decentralized Identifiers on Blockchains

More recently, with the advent of blockchain technology the notion of a resolvable identifier recorded on a blockchain has received considerable attention. Persistent identifiers which can be resolved to digital objects (e.g. files) have been in existence for nearly two decades now on the Internet. The most prominent of these identifier schemes is the *Digital Object Identifier* (DOI) [27], with its accompanying *Handle* resolver system [28], [29]. Similar in protocol-behavior to the DNS infrastructure, the DOI and Handle allows for efficient look-ups of copies of data files (e.g. library catalog entries) stored at open repositories all over the Internet. Currently the DOI/Handle system have been successfully deployed at a wide scale for over a decade. For example, most major academic journals today assign a persistent DOI identifier string to each published academic paper. This allows multiple copies of the paper to be stored at various repositories on the Internet, and any copy can be found by resolving DOI identifier string.

Today a new scheme being proposed – referred to *Decentralized Identifiers* (DID) [30] – extends this basic idea of the resolvable DOI by associating an identifier string belonging to a user with the public-key of the user and capturing this binding in a blockchain record. The DID structure is more generalized than the DOI/Handle in that it optionally allows the owner of the DID (e.g. the subject) to specify a service endpoint at which the caller may obtain corresponding services.

There are several benefits to this DID/blockchain approach: (i) only the user (possessing the matching private key) is able to create the DID record on the blockchain by digitally signing the transaction entry; (ii) only the user can "update" (add a new version) of the DID record to the blockchain; (iii) the DID record – and the identifier string in the record – is persistent (non-permutable) across time; and (iv) since each node in the blockchain network carries the full set of confirmed transactions, each node will carry a copy of all confirmed DID records, leading to higher availability.

Figure 3 summarizes the basic flow of information (claims) from a Claims Provider to the VASP. In Step (1) the subject (user) seeks the services of the VASP (e.g. transfer virtual asset). The subject must provide consent to the VASP by way of authorizing the Claims Provider in Step (2) to release the relevant claims to the VASP. The VASP then request access to the claims at the Claims Provider (e.g. by wielding an UMA authorization token) as shown in Step (3), with the Claims Provider responding by delivering the claims in Step (4).

An alternate flow using the DID/blockchain approach is shown in Figure 3 via Steps (3A) and (3B). Here the subject provides the VASP with a DID structure (either a public DID or pair-wise DID) in Step (1). The VASPs resolves the DID value (via the blockchain or DID resolver) in Step (3A), which brings the VASP to the correct Claims Provider – who holds the subject's claims – in Step (3B). As before the Claims Provider responds by delivering the signed claims in Step (4).

Although a detailed review of the DID protocol and methods are beyond the scope of the current work, there are several interesting aspects of the DID/blockchain approach.

First, the DID record (DID document) coupled with the blockchain capabilities has the potential to replace the classic Identity Provider (IdP) shown previously in Figure 2. In the classic IdP interaction, the subject (user) proves their identity by way of providing the IdP with their credential (e.g. password over SSL connection). In contrast, in the DID/blockchain approach the subject proves their identity by proving ownership control over the specific DID record on the blockchain. Since the DID record carries the subject's public key, any entity (e.g. a VASP) can easily obtain that public key and challenge the subject to sign a nonce or other challenge value. Secondly, the implications of this, among others, is that the subject (user) must now hold cryptographic keys instead of simply memorizing passwords. This further implies that some secure means must be used to hold, manage and exercise these private keys (e.g. secure mobile device with application software). This brings a whole range of security issues that are beyond the scope of discussion of the current paper.

Thirdly, the DID/blockchain approach in reality is unconnected (orthogonal) to the problem of the quality of claims or assertions. For VASPs and other Service Providers who rely on the truthfulness of claims or assertions issued by the Claims Provider, the degree of truthfulness (veracity) of any statement is crucial to their risk based decision-making and therefore to their business survival. The veracity of statements in a signed claim is not affected by the means of obtaining that claim (e.g. either directly from a Claims Provider, or through a DID/blockchain resolver redirection to the Claims Provider). More importantly, the quality of the claims produced by the Claims Provider is determined largely by the type of data and algorithms used by the Claims Provider (or Data Providers) [31] – which is a process that occurs external to the blockchain and the DID construct. The business risk of issuing claims is entirely burdened today by the Claims Provider and the subject.

Despite the recent hype about blockchain technology and "self sovereign" digital identity, the fact that a subject is reliant on other entities in society (e.g. banks, telecom providers, credit rating agencies) for sources of truthful information about the subject that affects their daily lives (e.g. for the purposes of obtaining goods and services, such as home loans) means that the subject is in reality not entirely self-sufficient ("self sovereign") [32], [33].

## V. Privacy-Preserving Claims: Open Algorithms

For Claims Providers and data holders generally, one important consideration for offering information (claims) about a subject is the impact on the subject in terms of data privacy and the fairness of the algorithms that are employed. This is because these aspects may impact the life of the subject (e.g. quality of credit score algorithm may impact ability to get a car loan). The notion of fairness applies both to the data used and to the algorithm design and construction [35], [36] Both privacy and fairness are complex issues that are beyond the current work.

We believe that alternative approaches to claims generation is needed that takes into account the fact that: (i) data is today siloed in various decentralized data repositories, (ii) that better insights can be derived when data from differing verticals or industry sectors are combined [36], and (iii) that data handling and privacy regulations may prohibit holders of data (referred to as *Data Providers*) from exporting it outside their organizational boundaries.

Our approach to decentralized data sharing is based on the *Open Algorithms* (OPAL) paradigm [37]. Briefly, the OPAL paradigm is founded on three basic principles. The first principle is that *data must never leave it repository*. Instead it is the algorithm that must be securely transmitted to the data repository and be executed there. This means that data must never be copied or duplicated from its repository. Second, only pre-approved algorithms that have been *vetted to be fair and unbiased* should be executed. Third, the default for responses (from algorithm executions) must be *aggregate-level responses* that preserves the privacy of the subjects whose data was involved in the algorithm computation. An aggregate-level response must be sufficiently granular that it prevent correlation attacks against the individual. Finally, a robust consent management protocol must be used which permits the logging and auditing of access to data and the executions of algorithms. Algorithms that are designed to identify individuals (e.g. who satisfy certain criteria) must be executed only after explicit consent has been obtain from the individual subject (following the GDPR [1]).

Thus, in order for VASPs to develop a CDD program that satisfies not only FinCEN and FATF requirements, but also preserves the privacy of citizens – as required by current privacy regulations (e.g. GDPR and CCPA) and possible future regulations [38] – we believe that the open algorithms paradigm offers a promising starting point to derive useful responses that can be conveyed in the claims format. A successful implementation of algorithms paradigm requires the participation of multiple Data Providers, arranged in a consortium-like organization we refer to as a *Trust Network* whose operations is governed by legally enforceable *System Rules*.

Figure 4 illustrates an OPAL-based trust network consists of a group of Data Providers who hold data about individual citizens and other legal entities. The consortium as a whole owns and operates a Claims Provider service, acing as a conduit for results of algorithmic computations conducted at the data repositories of the member Data Providers. Among others, the consortium must author and vet the algorithms that are permitted (accepted) to be executed on the repositories of the Data Providers. The trust network consortium may also operate an Authentication and Authorization Service (e.g. using UMA AS [24]) that implements consent management. It must also implement logging and audit mechanism in order to provider transparency to the trust network membership and achieve a higher degree of provenance regarding the insights or assertions that are encoded within the Claims format. Ideally, the Data Providers should consist of entities from differing industry sectors or vertical, such as telecom operators, banks, insurance providers, credit ratings entities, and so on.

Depending on the business model, the trust network consortium may require payment from external queriers (such as VASPs) for executing the algorithms on their data, and having their Claims Provider service issue a signed Claim. Queriers such as VASPs or other Service Providers should agree on the legal use terms of the signed Claims. Each signed Claim should be time-stamped and have a limited validity period (e.g. 24 hours, 7 days, 1 year) depending on the type of Claim and the use-case. Short validity periods discourages queriers from re-selling the signed Claims to the open market or to third party aggregators.

In Figure 4 before the VASP is permitted to engage the

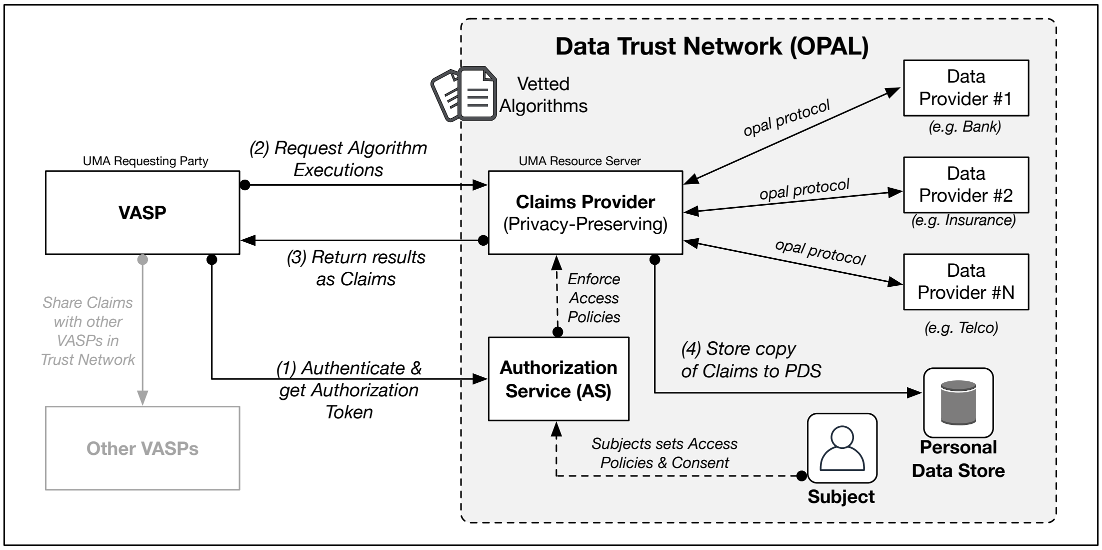


Fig. 4. The Data Provider Trust Network based on Open Algorithms (after [34])

Claims Provider service, the VASP as a relying party must first be authenticated and authorized by the *Authentication Service* (AS). This is shown in Step 1 of Figure 4. The VASP is permitted to choose only from a published list of vetted algorithms. In Step 2 the VASP submits a request to the Claims Provider. Responses coming back from the data providers are collated by the Claims Provider and packaged in the form of a claim or assertion using the relevant format (e.g. [18], [30]). The claims or assertions are digitally-signed by the Claims Provider, and then transmitted to the VASP in Step 3. A copy of all issued claims or assertions are also placed in the *claims store* of the subject located, for example, within the *Personal Data Store* (PDS) [39], [40] of the subject. The copies of signed claims in the subject's PDS claims-store allows the subject to independently make use of the claims for other purposes – which is consistent with the recommendation of the 2014 WEF report on personal data [8].

Although not shown explicitly in Figure 4, when a Data Provider executes a vetted algorithm and returns a response to the Claims Provider service in the data trust network, the Data Provider must include an indicator "score" relating to provenance and lineage of the data employed:

- *Provenance score*: This is an indicator of the degree of *creation origins* of the data. More specifically, this is an indicator if the data was generated by the institution (i) directly as part of their core business, (ii) as a byproduct of conducting business, (iii) imported into the institution directly from a known entity under a specific business agreement, (iv) imported into the institution from an uncertain origin (e.g. purchased from a 3rd party data aggregator).
  Thus, for example, a mobile telecom operator will generate Call Data Records (CDRs) and location data as a core part of doing business. This data is "directly generated" because they own the various physical networks, such as cell-towers, SS7 switching networks, routers, and others. Additionally, the mobile operator may also collect billing related data, including their customer's credit card information. This credit card information is considered a byproduct of doing business. (Note that a mobile operator corporation may own a consumer financing subdivision that may be a credit-card issuer. However, from a regulatory perspective this may be seen as a separate business for which current data flow limitations may apply).
- *Lineage score*: This is an indicator of the processing *generation* history of the data. Within many organizations, several steps of data processing (e.g. cleaning) may be applied. Additionally, specific versions of data-sets may be maintained by different business units.

## VI. The VASP Trust Network: Exchanging Claims and Key Information

As mentioned in Section V, using the open algorithms paradigm VASPs are able to obtain privacy-preserving claims from the community of Data Providers whose constituents may come from different market verticals (Figure 4). Additionally, by interacting with the community of certificate authorities (CAs) a VASP may also obtain assurance of the correct legal ownership of private-public keys. From the Data Providers perspective, virtual assets and VASPs represents a new market and new revenue source for the privacy-safe use of data in their possesion.

A current issue for many large data-holding institutions is knowing the origins of data in their possession (e.g. in their "data lakes") and the subsequent processes applied to the data (i.e. ETL processing) as it is employed within different parts of the organization [41]. For usage in Claims, the data provenance and lineage score represents valuable information to the VASPs for CDD requirements. Figure 5 illustrates a simple VASP trust network, where each VASP obtains Claims

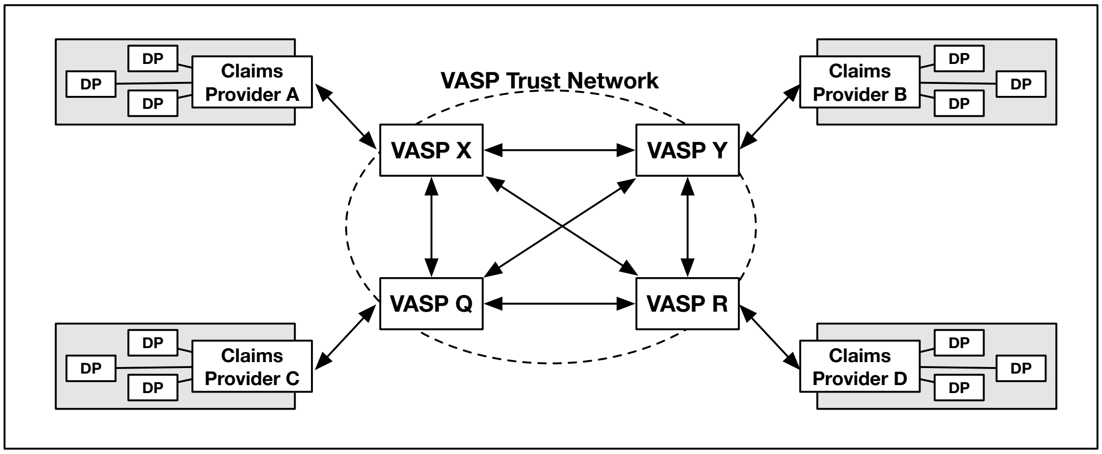


Fig. 5. The Trust Network of VASPs for Claims Exchanges

from their respective Claims Provider (each of which belongs to a different OPAL-based Data Trust Network).

### A. VASP Exchanges of Claims

In order to make an OPAL-based Claim to be accurate and useful for VASPs and other claims-consumers (e.g. traditional FIs), there are a number of rudimentary metadata information that needs to be captured and represented in a Claim structure:

- *Serial number of Claim*: This is the globally unique serial number os this claim, which may consists of a simple hash of other combined fields.
- *Identity of issuing Claims Provider*: This is the identifier of the Claims Provider (or Claims Issuer) that signs the Claim structure
- *Algorithm identifier employed to create the Claim*: This is a set of identifier strings which have local meaning among the Data Provider trust network.
- *Provenance & lineage scores*: The Claims Provider needs to communicate some measure of confidence in the accuracy of the statement contained in the Claim.
- *Hashes of data-sets identifiers*: Optionally, Data Provider may record the internal data-sets used in the computation by way of computing hashes upon the data-set (assuming its is not CPU intensive) and use this hash value as an identifier within the Data Provider trust network.

It is worth nothing that there are already several Claims format structures developed over the past two decades (e.g. X.509 Attribute certificates [42], SAML assertions [18], Verifiable Claims [30]).

### B. VASP Exchanges of Key Information

VASPs should exchange either serial-number of certificates or hash-values of public-keys among each other in the trust network. Similarly, the trust network of VASPs should periodically (e.g. hourly, overnight) exchange the certificate revocation list (CRL) [43], [44] among the members of the trust network using the existing X.509 standard protocols.

This list of known good serial numbers and public keys could be shared (broadcasted) securely with members of the trust network on a regular basis (e.g. hourly or overnight) based on a push/pull model (e.g. over a RESTful API [45]). This allows one VASP to query another VASP in the trust network, submitting only the serial numbers or the public keys over a point-to-point secure channel (e.g. SSL).

## VII. Conclusions

Currently in order for VASPs to fulfill the regulatory requirements from the FATF and the Travel Rule, VASPs need access to truthful information – or claims – regarding originators, beneficiaries and other VASPs involved in a virtual asset transfer instance. Obtaining access to data or information regarding individuals and institutions brings with it another set if challenges relating to data privacy and privacy-related regulations.

For VASPs to develop a CDD program that satisfies not only FinCEN and FATF requirements, but also preserves the privacy of citizens as required by current privacy regulations (e.g. GDPR and CCPA), we believe that the open algorithms paradigm offers a promising starting point to derive useful responses that can be conveyed in the claims format. A successful implementation of the open algorithms paradigm requires the participation of multiple data providers, something which can be achieved using a shared trust network for claims exchange that is governed by a common legal trust framework defined by the VASP communities globally.

## References

[1] European Commission, "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)," *Official Journal of the European Union*, vol. L119, pp. 1–88, 2016.

[2] California State Legislature, "California Consumer Privacy Act (CCPA) - AB 375," California Civil Code – Section 1798.100, September 2018.

[3] G. Chavez-Dreyfuss, "Cryptocurrency theft hits nearly $1 billion in first nine months," *Reuters Business News*, October 2018. [Online]. Available: https://www.reuters.com/article/us-crypto-currency-crime/cryptocurrency-theft-hits-nearly-1-billion-in-first-nine-months-report-idUSKCN1MK1J2

[4] D. Shane, "$530 million cryptocurrency heist may be biggest ever," *CNN Business*, January 2018. [Online]. Available: https://money.cnn.com/2018/01/29/technology/coincheck-cryptocurrency-exchange-hack-japan/index.html

[5] FATF, "International Standards on Combating Money Laundering and the Financing of Terrorism and Proliferation," Financial Action Task Force (FATF), FATF Revision of Recommendation 15, October 2018, available at: http://www.fatf-gafi.org/publications/fatfrecommendations/documents/fatf-recommendations.html.

[6] Financial Crimes Enforcement Network (FinCEN) - Department of the Treasury, "Customer Due Diligence Requirements for Financial Institutions (31 CFR Parts 1010, 1020, 1023, 1024, and 1026; RIN 1506AB25)," *Federal Register*, vol. 79, no. 149, August 2014, available at: https://www.fincen.gov/sites/default/files/shared/CDD-NPRM-Final.pdf.

[7] World Economic Forum, "Personal Data: The Emergence of a New Asset Class," 2011, http://www.weforum.org/reports/ personal-data-emergence-new-asset-class.

[8] ——, "Rethinking Personal Data: A New Lens for Strengthening Trust," May 2014, http://reports.weforum.org/rethinking-personal-data.

[9] M. Madden, "Public Perceptions of Privacy and Security in the Post-Snowden Era," November 2014, http://www.pewinternet.org/2014/11/12/public-privacy-perceptions/.

[10] R. Abelson and M. Goldstein, "Millions of Anthem customers targeted in cyberattack," *New York Times*, February 2015. [Online]. Available: https://www.nytimes.com/2015/02/05/business/hackers-breached-data-of-millions-insurer-says.html

[11] T. S. Bernard, T. Hsu, N. Perlroth, and R. Lieber, "Equifax says cyberattack may have affected 143 million in the U.S." *New York Times*, September 2017. [Online]. Available: https://www.nytimes.com/2017/09/07/business/equifax-cyberattack.html

[12] FATF, "Financial Action Task Force (FATF)," 2019, http://www.fatf-gafi.org.

[13] ——, "Guidance for a Risk-Based Approach to Virtual Assets and Virtual Asset Service Providers," Financial Action Task Force (FATF), FATF Guidance, June 2019, available at: www.fatf-gafi.org/publications/fatfrecommendations/documents/Guidance-RBA-virtual-assets.html.

[14] T. Hardjono, "Compliant Solutions for VASPs," May 2019, presentation to the FATF Private Sector Consultative Forum (PSCF) 2019, Vienna (6 May 2019).

[15] T. Hardjono, A. Lipton, and A. Pentland, "Towards a Public Key Management Framework for Virtual Assets and Virtual Asset Service Providers," 2020, Journal of FinTech (to appear) – Available at https://arxiv.org/pdf/1909.08607.

[16] W. Simpson, "Ppp challenge handshake authentication protocol (chap)," August 1996, RFC1994. [Online]. Available: http://tools.ietf.org/rfc/rfc1994.txt

[17] T. Hardjono, "Federated Authorization over Access to Personal Data for Decentralized Identity Management," 2019, IEEE Communications Magazine (to appear) – Available at https://arxiv.org/pdf/1906.03552.pdf.

[18] OASIS, "Assertions and Protocols for the OASIS Security Assertion Markup Language (SAML) V2.0," March 2005, available on http://docs.oasisopen.org/security/ saml/v2.0/ saml-core-2.0-os.pdf.

[19] American Bar Association, "An Overview of Identity Management: Submission for UNCITRAL Commission 45th Session," ABA Identity Management Legal Task Force, May 2012, available on http://meetings.abanet.org/ webupload/commupload/ CL320041/relatedresources/ ABA-Submission-to-UNCITRAL.pdf.

[20] OASIS, "Glossary for the OASIS Security Assertion Markup Language (SAML) V2.0," March 2005, available on http://docs.oasis-open.org/security/ saml/v2.0/samlglossary- 2.0-os.pdf.

[21] Gartner, "2017 Planning Guide for Identity and Access Management," Gartner Inc., Report, October 2016.

[22] K. Cameron, "The Laws of Identity," 2005. [Online]. Available: https://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf

[23] T. Hardjono, E. Maler, M. Machulak, and D. Catalano, "User-Managed Access (UMA) Profile of OAuth2.0 – Specification Version 1.0," Kantara Initiative, Kantara Published Specification, April 2015, https://docs.kantarainitiative.org/uma/rec-uma-core.html.

[24] E. Maler, M. Machulak, and J. Richer, "User-Managed Access (UMA) 2.0," Kantara Initiative, Kantara Published Specification, January 2017, https://docs.kantarainitiative.org/uma/ed/uma-core-2.0-10.html.

[25] M. Lizar and D. Turner, "Consent Receipt Specification Version 1.0," March 2017, https://kantarainitiative.org/confluence/display/infosharing/Home.

[26] D. Hardt, "The OAuth 2.0 Authorization Framework," October 2012, RFC6749. [Online]. Available: http://tools.ietf.org/rfc/rfc6749.txt

[27] ISO, "Digital Object Identifier System – Information and Documentation," International Organization for Standardization, ISO 26324:2012, June 2012, available at: http://www.iso.org/iso/catalogue_detail?csnumber=43506.

[28] S. Sun, L. Lannom, and B. Boesch, "Handle System Overview," November 2003, RFC3650. [Online]. Available: http://tools.ietf.org/rfc/rfc3650.txt

[29] S. Sun, S. Reilly, and L. Lannom, "Handle System Namespace and Service Definition," November 2003, RFC3651. [Online]. Available: http://tools.ietf.org/rfc/rfc3651.txt

[30] M. Sporny, D. Longley, and D. Chadwick, "Verifiable Credentials Data Model 1.0," W3C, W3C Recommendation, November 2019, available at https://www.w3.org/TR/verifiable-claims-data-model.

[31] T. Hardjono and A. Pentland, "Core Identities," in *Trusted Data - A New Framework for Identity and Data Sharing*, T. Hardjono, A. Pentland, and D. Shrier, Eds. MIT Press, 2019, pp. 41–81.

[32] A. Pentland and T. Hardjono, "Digital Identity Is Broken. Heres a Way to Fix It," *Wall Street Journal*, April 2018, https://blogs.wsj.com/cio/2018/04/03/ digitalidentityisbrokenhere-sawaytofixit/.

[33] T. Hardjono and E. Maler, "Blockchain and Smart Contracts Report," Kantara Initiative, Report, June 2017, https://kantarainitiative.org/confluence/display/BSC/Home.

[34] T. Hardjono and A. Pentland, "Open Algorithms for Identity Federation," in *Advances in Information and Communication Networks: Proceedings of the 2018 Future of Information and Communication Conference (FICC), Vol. 2*, K. Arai, S. Kapoor, and R. Bhatia, Eds. Springer-Verlag, 2018, pp. 24–43.

[35] M. A. Bakker, D. P. Tu, H. R. Valds, K. P. Gummadi, K. R. Varshney, A. Weller, and A. Pentland, "DADI: Dynamic Discovery of Fair Information with Adversarial Reinforcement Learning," 2019, Available at: https://arxiv.org/abs/1910.13983.

[36] A. Pentland, *Social Physics: How Social Networks Can Make Us Smarter*. Penguin Books, 2015.

[37] T. Hardjono and A. Pentland, "MIT Open Algorithms," in *Trusted Data - A New Framework for Identity and Data Sharing*, T. Hardjono, A. Pentland, and D. Shrier, Eds. MIT Press, 2019, pp. 83–107.

[38] C. F. Kerry, "A federal privacy law could do better than Californias," Brookings Institution, Report - Center for Technology Innovation, April 2019, https://www.brookings.edu/blog/techtank/2019/04/29/a-federal-privacy-law-could-do-better-than-californias/.

[39] T. Hardjono and J. Seberry, "Strongboxes for Electronic Commerce," in *Proceedings of the Second USENIX Workshop on Electronic Commerce*. Berkeley, CA, USA: USENIX Association, 1996.

[40] Y. A. de Montjoye, E. Shmueli, S. Wang, and A. Pentland, "openPDS: Protecting the Privacy of Metadata through SafeAnswers," *PLoS ONE 9(7)*, pp. 13–18, July 2014, https://doi.org/10.1371/journal.pone.0098790.

[41] N. Heudecker, E. Thoo, and E. Zaidi, "Critical Capabilities for Data Integration Tools," Gartner, Research Report G00377592, October 2019.

[42] S. Farrell, R. Housley, and S. Turner, "An internet attribute certificate profile for authorization," January 2010, RFC5755. [Online]. Available: http://tools.ietf.org/rfc/rfc5755.txt

[43] R. Housley, W. Polk, W. Ford, and D. Solo, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile," April 2002, RFC3280. [Online]. Available: http://tools.ietf.org/rfc/rfc3280.txt

[44] D. Cooper, S. Santesson, S. Farrell, S. Boeyen, R. Housley, and W. Polk, "Internet X.509 public key infrastructure certificate and certificate revocation list (crl) profile," May 2008, RFC5280. [Online]. Available: http://tools.ietf.org/rfc/rfc5280.txt

[45] P. Gutmann, "Internet X.509 public key infrastructure operational protocols: Certificate store access via http," February 2006, RFC4387. [Online]. Available: http://tools.ietf.org/rfc/rfc4387.txt