Blockchain-based Identity Management and Distribution

2.1 Introductions

Everyday, it seems, media reports about various services based on blockchain technology appear. Among these are many unfortunate ideas that merely use blockchains as a distributed database, so much so that a number of flowcharts for determining whether you really do need to use blockchains have been published[*1]. There are several methods of classifying blockchains; broadly, they can be classified into private-use blockchains and public blockchains that underpin the security of cryptographic assets. With public blockchains, it is necessary to incentivize ongoing mining to extend the chain; for cryptoassets such as Bitcoin, it is necessary to extend the chain based on predefined rules. With cryptoassets, the main reason for using blockchain is to transfer assets from one address to another, but efforts are also being made to use this blockchain-based value-transfer platform for other applications. These are being called second layer or Layer 2 applications.

Here, we look at developments in ERCs (Ethereum Requests for Comment)[*2] used for credentials (identity information) from among the second-layer services for the Ethereum blockchain[*3]. We will also touch on use cases in which such credentials are stored in the blockchain, enabling public certifications, such as student or employee IDs, to be verified digitally. Finally, we also go over why several vendors and consortiums have put forward concepts such as Decentralized Identifiers (DIDs) and Self-Sovereign Identity (SSI), where the user is in control of managing their own identity, and take a look at the focus on technology for managing credentials based on blockchains.

2.2 IDs and Credentials as Identifiers

In focused research pieces back in 2015, we reported on trends in ID management technology at the time[*4*5*6]. Here, we consider IDs in the narrow sense of identifiers.

Real-world entities are linked with digital-world entities, and a unique identifier (which we will denote "ID") is assigned to identify the digital-world entity. The notion of an ID as an identifier must be kept conceptually separate from the various identity information that is bound to that ID. Further, because realms (the scope within which the ID is valid and can be used to identify something) are separately defined for each ID space, a single, unique entity in the real world can have multiple IDs even within the same realm.

Now the reason IDs are assigned in the digital space is that there is a need for third parties on the network to identify the entities to which IDs are assigned. The act of authentication accompanies all sorts of activities in the digital world. Authentication allows access to various resources and services, for instance.

This act of authentication can be accomplished by using pairs of tokens (secret information) and credentials (public information). According to the definition in NIST SP 800-63, a token contains a secret known to the user to which the ID in question has been assigned, and credentials bind various attributes to the ID. Cryptographic techniques are used to ensure the integrity of credentials. Credentials use cryptographic techniques to ensure content integrity. When the entity that holds the ID seeks verification of his or her attributes by a third party in the digital world, the token (secret information) can be used to verify that the holder of the ID is the entity to which it was assigned.

When a credential is presented together with an authentication operation, a receiving third party can verify what sort of entity the ID is using the attributes given in the credential. In addition to authentication in this manner, credentials are also used for authorization in some cases. An X.509 certificate is

*1 NISTIR 8202, "Blockchain Technology Overview" (https://doi.org/10.6028/NIST.IR.8202), Figure 6: DHS Science & Technology Directorate Flowchart.

*2 Ethereum Improvement Proposals (http://eips.ethereum.org/).

*3 Ethereum Project, Developer Resources (https://www.ethereum.org/developers/). In this volume, we do not cover technology related to smart contracts, a key feature of Ethereum.

*4 Internet Infrastructure Review Vol. 26, "1.4.3 ID Management Technology" (https://www.iij.ad.jp/en/dev/iir/pdf/iir_vol26_EN.pdf).

*5 Internet Infrastructure Review Vol. 27, "1.4.2 ID Management Technology: From a Convenience and Security Perspective" (https://www.iij.ad.jp/en/dev/iir/pdf/iir_vol27_EN.pdf).

*6 Internet Infrastructure Review Vol. 28, "1.4.3 ID Management Technology: Online Authentication Methods Not Using Passwords" (https://www.iij.ad.jp/en/dev/iir/pdf/iir_vol28_EN.pdf).

an example of a credential because it binds a public key with one or more IDs. And in fact, SSL/TLS client authentication is one case of this. Deploying a personal X.509 certificate on the browser side allows a user to log in to a server, and this is used in applications like corporate online banking. A specification for X.509 Attribute Certificates[*7] provides a method that is closer to credential-like usage. Attribute certificates differ from ordinary X.509 certificates in that they do not contain a public key. A serial number used to identify the certificate is placed in an area for storing identifiers called the Holder so as to specify the X.509 certificate, and attributes bound to the certificate holder (subject) are then stored. Here the realm can be understood to be the certificates issued by the certification authority, the ID to be the serial number, and the credentials to be the attribute certificate. Credentials can be written to X.509 Attribute Certificates, but they are not actually implemented in applications that general users are likely to encounter, such as browsers, so there are hardly any cases of them being used at present.

2.3 Overview of ERC-725

ERC-725[*8] was proposed in October 2017 by software engineer Fabian Vogelsteller[*9], known for developing the ERC-20 token standard and web3.js. Like the IETF's RFCs, ERCs (Ethereum Requests for Comment) are documented proposals for improvements that anyone can author; the format and writing guidelines are given in ERC-1. A major feature worth noting is that authors are asked to keep their proposals compact. A similar class of documents exists in the Bitcoin community, known as BIPs (Bitcoin Improvement Proposals)[*10]. The method for reducing transaction data known as SegWit, for instance, is defined in BIP-141.

Smart contracts, a method for automatically executing contracts and performing services, are a new concept put forward by Nick Szabo in 1997 and thus predate Bitcoin. A commonly cited example of a smart contract is the vending machine. When certain conditions are met by two processes, namely that the user inserts payment for the desired beverage into the machine and the user subsequently presses the button corresponding to that beverage, a sale is automatically initiated. As well as being used for cryptoassets, Ethereum is also being viewed as a distributed application platform enabling the creation and execution of smart contracts[*11]. ERC-725 defines a Solidity interface for the behavior of a proxy account. Solidity is a language used to write distributed applications. ERC-725 refers to ERC-735[*12] and ERC-780[*13] and provides a framework for distributing credentials on the Ethereum blockchain based on these specifications. In the ERC-725 document, credentials are called claims. ERC-735 describes the format of claims, and ERC-780 describes an Ethereum Claims Registry (ECR). Under the framework specified in the Ethereum blockchain realm, an ID (identifier) is an Ethereum address (note that it is not a contract address), and the identity information of the identity holder bound to the address is certified by credentials, called claims. The claim issuer can issue a claim to any entity on the Ethereum blockchain using a private key in the claim issuer's possession. The identity holder passes the claim to be verified to the claim checker via some method, and the claim checker can verify the claim's veracity by verifying the digital signature. It is envisioned that this series of verification tasks can be performed both online and offline[*14].

Claims as specified in ERC-735 have the following simple data structure.

*7 RFC 5755, "An Internet Attribute Certificate Profile for Authorization" (https://datatracker.ietf.org/doc/rfc5755/).
*8 ERC-725 version 2: Proxy Account (https://github.com/ethereum/EIPs/issues/725) (http://eips.ethereum.org/EIPS/eip-725).
*9 Fabian Vogelsteller (http://frozeman.de/blog/).
*10 BIP (Bitcoin Improvement Proposals) (https://github.com/bitcoin/bips).
*11 Ethereum Project white paper (https://github.com/ethereum/wiki/wiki/White-Paper).
*12 ERC-735: Claim Holder (https://github.com/ethereum/EIPs/issues/735).
*13 ERC-780: Ethereum Claims Registry (https://github.com/ethereum/EIPs/issues/780).
*14 Fabian Vogelsteller, ERC Identity (https://www.slideshare.net/FabianVogelsteller/erc-725-identity).

```
struct Claim {
    uint256 topic;
    uint256 scheme;
    address issuer;
    bytes signature;
    bytes data;
    string uri;
}
```

ERC-735 claims should be implemented to enable the identity holder to present them to the claim checker, and a key characteristic is that the data are portable. ERC-735 provides a zone for writing URIs to an area that is not ToBeSigned, and it is here that data pointing to the identity information is shared via a distributed file system such as IPFS[*15]. The ERC-725 Alliance[*16] has an open-source project related to ERC-725[*17]. Also, a number of samples are available on sites[*18] built using this demo implementation, showing how the veracity of claims can be verified in the browser. It is worth noting that the specification allows you to sign your own identity information and thus make your own claims about yourself.

So under the ERC-725 framework, anyone can issue a claim. In other words, anyone can be a claim issuer, so you can issue a claim to anyone as long as you know their Ethereum address. A key issue, therefore, is how to establish trust for a claim issuer and how to value the claims issued by that issuer. There also appears to be functionality to allow, for example, claims to be revoked and the Ethereum address to be swapped out, but the specification is still incomplete in this regard. It also seems that discussion over what reputation system to use for issuers has only just begun.

So we are straddled with a problem of reputation, and we will probably need to work through a few stages before we are ultimately able to distribute claims the way we would like. My view is that the notion of claims will gradually gain traction via the following three steps. In the first stage, acquaintances in closed networks, such as SNS, will casually issue claims to one another. This phase will determine scalability. The next stage will see the formation of a framework

Table 1: Elements of the ERC-735 Claim Structure

Topic	Currently marked as ToBeDefined. A 256-bit space expected to contain information on the topic (or type) of claim.
Scheme	ToBeDefined. A 256-bit space to hold the processing method or signature algorithm to use, which would refer to separately defined schemes.
Issuer	A contract address or the Ethereum address of the key used to sign the signature.
Signature	Note that the signed data needs to be of the following structure: {identity holder's Ethereum address, topic, data}
Data	The hash of the identity information (claim data). The identity information itself is not written here, so sensitive information is not being written to the blockchain.
Uri	A URI pointing to the identity information. HTTP link, IPFS URI, or such like.

*15 IPFS (InterPlanetary File System) (https://docs.ipfs.io/introduction/).

*16 ERC-725 Alliance (https://erc725alliance.org/).

*17 ERC-725 Alliance, "Repository for code and discussion around ERC725 and related standards" (https://github.com/ERC725Alliance/erc725/tree/master/contracts/contracts).

*18 ERC 725: Demo implementation by Origin Protocol (https://erc725.originprotocol.com/); Origin Protocol, Inc. (https://www.originprotocol.com/).

for ranking issuers using existing user evaluation/reporting systems to assess whether they have issued incorrect claims or not. And finally, this will develop into a completely distributed, automated reputation system (Figure 1).

The Blockcerts[*19] project is another example of the use of portable claims in a vein similar to ERC-725; there are open-source code[*20] and verification demos[*21] available as well. Blockcerts is based on prototypes developed at the MIT Media Lab and Learning Machine. Work is ongoing to extend Blockcerts to implement multiple blockchain types, including Bitcoin and Ethereum. A smartphone app called Blockcerts Wallet has also been implemented and released, and MIT is now using the Blockcerts technology to write students' diplomas to the blockchain[*22]. And a Spanish university has also announced that when issuing degree certificates, it will use the SmartDegrees platform so that they can be managed on the Ethereum blockchain[*23]. The situation is a tad chaotic at present with multiple such second-layer platforms on the scene, so when selecting a platform, the business continuity prospects need to be taken into account.

As discussed so far here, claims embody a simple mechanism, but if the Issuer is trustworthy and the second-layer specification works properly, it is understood they can be used on a semipermanent basis so long as the reliability of the Ethereum blockchain remains intact. So the digital issuance of diplomas is one application apt for making good use of blockchain technology, and indeed, some such services have appeared in Japan. Once trusted organizations do not persist indefinitely, as attested by the closure of private schools in regional areas and the discontinuation of certifying exams by local governments. There are even cases of physical certificates issued by such organizations no longer being validatable. Hopefully, we are bound for an era in which claims distributed via an open framework, as discussed here, provide an alternative to physical certificates.

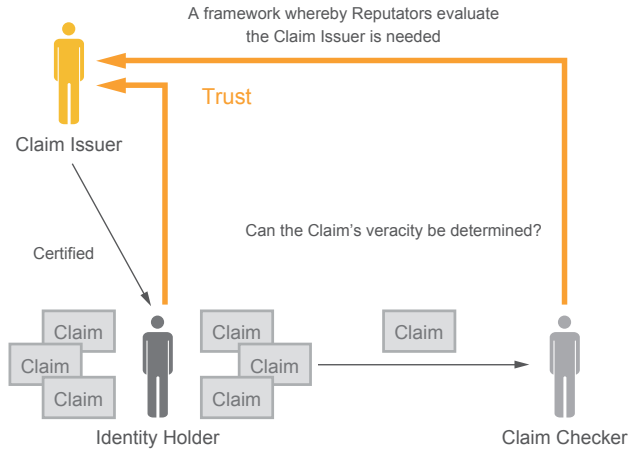

Figure 1: Framework for Issuing and Valuing Claims

*19 Blockcerts (https://www.blockcerts.org/).
*20 Repositories of the Blockcerts project (https://github.com/blockchain-certificates).
*21 Example Blockcerts (https://www.learningmachine.com/new-product/examples/).
*22 MIT News, Digital Diploma debuts at MIT (http://news.mit.edu/2017/mit-debuts-secure-digital-diploma-using-bitcoin-blockchain-technology-1017).
*23 Universidad Carlos III de Madrid is issuing degree certificates with blockchain (https://www.uc3m.es/ss/Satellite/UC3MInstitucional/en/Detalle/Comunicacion_C/1371252827656/1371215537949/Universidad_Carlos_III_de_Madrid_is_issuing_degree_certificates_with_blockchain).

2.4 Decentralized Identifiers, DIDs

As identifiers, IDs are assigned within a specific realm. When it comes to authenticating across realm boundaries, the notion of the ID Federation comes into play and often appears in a single sign-on context. Credentials such as the X.509 Attribute Certificates and ERC-735 claims we have discussed only circulate within the realm in which they were issued. In reality, identity providers, whose role is also to issue IDs, do not exist in isolation. To enable the login functionality of service providers, such as SNSs, to be used from external services, that functionality is split off into the identity provider role. As such, in cases where ID linking functionality is used to log in to separate services, there is a risk that the ID will suddenly stop working because it is operated by a particular company or organization. Thus, the deactivation of one ID could result in an inability to use several other services. The deactivation—or in the worst case, deletion—of an ID because an SNS operator decided that inappropriate content had been posted can have negative impacts, and indeed there have been various real-world cases of this.

The notion of Decentralized Identifiers (DIDs) emerges from this backdrop. A feature of DIDs is that they are not valid IDs only in one specific realm and there is no centralized presence that manages the IDs. This is seen as highly compatible with the notion of Self-Sovereign Identity (SSI)[*24] proposed by the nonprofit Sovrin Foundation[*25]. SSI is similar to the idea that individuals have the right to control their own information. The term is used in recognition of the need for individuals to own and manage their own identities without going through a central managing authority. Credentials such as ERC-735 claims, as discussed above, can be passed around without the identity holder intending it. Not so with SSI; instead, the idea is that the identity holder has sovereignty over the distribution of his or her credentials.

The nonprofit ID2020 Alliance[*26] is an organization that seeks to achieve privacy protection and portable, user-centric identity management. There is also a project[*27] that looks to use claims written to a blockchain as an alternative to passports, analogous to the way people seek to use

*24 The Sovrin Alliance, "Sovrin: A Protocol and Token for Self-Sovereign Identity and Decentralized Trust" (white paper, version 1.0, January 2018) (https://sovrin.org/wp-content/uploads/Sovrin-Protocol-and-Token-White-Paper.pdf).

*25 The Sovrin Alliance (https://sovrin.org/library/rise-of-self-sovereign-identity/).

*26 ID2020 Alliance, The Alliance Manifesto (https://id2020.org/manifesto).

*27 Taqanu (https://www.taqanu.com/impact).

cryptoassets rather than legal currency in cases where the reliability of nationally issued currencies has diminished. The idea can be interpreted as follows: data corresponding to claims that everyone recognizes and that are issued by credible institutions can provide a passport-like means of personal identification. The Sustainable Development Goals (SDGs)[*28] compiled by the United Nations in 2015 state Target 16.9 as: "By 2030, provide legal identity for all, including birth registration". And the ID2020 Technical Requirements[*29] have been formulated in an attempt to assist the world's "identity refugees", thought to number over a billion. The document covers seven categories—applicability, identification, authentication, privacy, trust, interoperability, and recovery—and is highly useful as a design guideline of this type.

The intention behind DIDs, meanwhile, can be gleaned from documents developed by the W3C[*30][*31]. The W3C defines a DID as a globally unique identifier that does not require a centralized registration authority because it is registered with distributed ledger technology or other form of decentralized network[*32]. ERC-735 claims use Ethereum addresses as the ID space, but a method has also been proposed for wrapping Ethereum addresses in the W3C DID format instead of using them as raw DIDs[*33]. So W3C DID is being promoted as a global ID capable of representing a range of IDs. The existence of DIDs alone only solves the issue of nonconflicting numbering, but in conjunction with the claim use cases[*34] and the verifiable credentials (originally called claims, the wording was later changed to credentials) data format[*35], they are poised to solve the various other issues faced.

Group work at the Web of Trust VIII event in March 2019 (RWOT8)[*36] and the 28th Internet Identity Workshop[*37] in April 2019 dealt with many topics centering on DIDs and SSI. The 2019 annual meeting of the IGF (Internet Governance Forum)[*38] will also cover DID-related technology and encompass discussion of governance. So looking ahead, many people are lining up to drive the discussion forward.

*28 Transforming our world: the 2030 Agenda for Sustainable Development (https://sustainabledevelopment.un.org/post2015/transformingourworld) (https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E).
*29 ID2020 Technical Requirements: V1.0 (https://docs.google.com/document/d/1L0RhDq98xj4ieh5CuN-P3XerK6umKRTPWMS8Ckz6_J8/edit).
*30 W3C Credentials Community Group (https://www.w3.org/community/credentials/).
*31 W3C Verifiable Claims Working Group (https://www.w3.org/2017/vc/WG/) (https://github.com/w3c/verifiable-claims).
*32 Decentralized Identifiers (DIDs) (https://w3c-ccg.github.io/did-spec/#decentralized-identifiers-dids); latest version as of this writing: v0.13, dated Jun. 3, 2019
*33 eth DID Resolver (https://github.com/uport-project/eth-did-resolver).
*34 Verifiable Claims Use Cases (https://www.w3.org/TR/verifiable-claims-use-cases/).
*35 Verifiable Credentials Data Model 1.0 (https://www.w3.org/TR/verifiable-claims-data-model/); latest version: Mar. 2019; a W3C Candidate Recommendation as of this writing.
*36 Rebooting the Web of Trust VIII: Barcelona (March 2019) (https://github.com/WebOfTrustInfo/rwot8-barcelona) (https://www.weboftrust. info/pastevents.html).
*37 IIW (The Internet Identity Workshop) Workshop Proceedings (https://internetidentityworkshop.com/past-workshops/).
*38 IGF 2019 Workshop Selection Results (https://www.intgovforum.org/multilingual/content/igf-2019-workshop-selection-results).

2.5 Other Related Developments

In May 2019, Microsoft unveiled a platform to handle DIDs based on the Bitcoin blockchain. Two blog posts on May 15 describe its future activities in this area[*39][*40]. And it has published a white paper on DID[*41]. From these sources, it is apparent that the W3C DID is being used as the ID space and that the Sidetree protocol developed by the Decentralized Identity Foundation (DIF)[*42] has been adopted. This DID system is implemented on the second layer of the Bitcoin blockchain, and source code has already been released under the name ION (Identity Overlay Network)[*43].

Finally, I will touch on credit scores and information banks. Some media reports claim services that calculate credit scores based on online activity are in the works in Japan as well. A concern is that only scoring done under the auspices of big-brother entities like GAFA and FAANG would be considered accurate, and that your score could be passed around without you intending it. As with AI, another talking point of late, there is a need to ensure transparency of not only the scoring system but the scoring algorithm as well. It is theoretically possible for people to be scored unfairly based on obscure logic because they live in a particular region or on the basis of race, religion, etc. Hence, ethical considerations must be taken into account. The same can be said for the reputation mechanisms of real-world entities, a concern I also noted in relation to ERC-735 claims.

Thus, we now find ourselves in an age in which real-world entities are subject to being evaluated by various means. From a management perspective, these measures may be necessary to ensure security, yet we still need a way to enable individuals to manage their own sensitive information based on the SSI concept. In particular, although it may not be easy to ask identity holders who have been issued a DID or claim for the first time to protect themselves in the ways required, I think the ability to do this really is part of the basic literacy we all need as denizens of the digital age.

Some people have a desire to pass their social media accounts to their children or grandchildren after they die, but this is becoming less and less advisable from a business

*39 Microsoft Security Blog, "Decentralized identity and the path to digital privacy" (https://www.microsoft.com/security/blog/2019/05/15/decentralized-identity-digital-privacy/).

*40 Azure Active Directory Identity Blog, "Toward scalable decentralized identifier systems" (https://techcommunity.microsoft.com/t5/AzureActive-Directory-Identity/Toward-scalable-decentralized-identifier-systems/ba-p/560168).

*41 Microsoft Whitepaper, Decentralized Identity—Own and control your identity (https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE2DjfY).

*42 Decentralized Identity Foundation (DIF) (https://identity.foundation/) (https://github.com/decentralized-identity/).

*43 ION (Identity Overlay Network) (https://github.com/decentralized-identity/ion/).

continuity perspective. With the advance of AI, it seems, the temporary or permanent deactivation of accounts as an act of censorship against posted content is having a major impact. I think these sorts of "everyday" examples are also a factor behind the rising call for services based on DID and related technologies.

At present, I think a lot remains to be discussed in regard to how we treat temporary IDs when it comes to handling massive quantities of statistical information and with respect to cases in which credentials themselves are encrypted as part of access control. A technology does not necessarily gain traction just because the background technologies are compatible and it would have social applicability if deployed adroitly, and I have seen this many times over the years. At this juncture, it is uncertain whether the technologies I have discussed here will be deployed in applications users identify with and be of any use to the world.

With the advent of real use cases such as information banks, people have come to realize the convenience provided by mechanisms for the Internet-based distribution of information (including sensitive information) linked to real-world entities. But we face a large problem here. Privacy regulations like the EU's GDPR are not unique to the EU sphere. Countries around the world, including Japan, are also subject to such regulations. Hence, because the technologies discussed here use blockchain and the circulation of credentials is thus not limited to within any one region, use of such technologies could face restrictions according to the range of regulations that countries around the world have in place. This is a far cry from the thinking behind cryptoassets such as Bitcoin and Ethereum, and could be a major factor impeding the penetration of such technologies. The ERC-725 Alliance and ID2020 will need to undertake activities to dispel these impediments to cross-country distribution mechanisms. At present, though, no such activities appear to be underway. We need experts that can offer deep insights and broad perspectives to address these issues, including the issue of whether such problems should be dealt with by these consortiums in the first place.

Yuji Suga

Senior Engineer, Office of Emergency Response and Clearinghouse for Security Information, Advanced Security Division, IIJ. Dr. Suga has been in his current position since July 2008. He is engaged in investigation and research activities related to cryptography and information security as a whole. He heads up the CRYPTREC Operational Guidelines Working Group on TLS Configuration and is a member of the CRYPTREC Cryptographic Technology Promotion Committee. He also serves as secretariat of the Cryptographic protocol Evaluation toward Long-Lived Outstanding Security Consortium (CELLOS); secretary of the Information Processing Society of Japan's Computer Security Group (CSEC); assistant secretary of the ISEC Technical Committee of the Institute of Electronics, Information and Communication Engineers (IEICE); CyberSciTech2019 program co-chair; organizing committee member for IWSEC2019; and member of the Cryptoassets Governance Task Force (CGTF) Security Working Group.